PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/82, 9/00, A01H 5/00, C07H 21/02	A2	(11) International Publication Number: **WO 99/28485** (43) International Publication Date: 10 June 1999 (10.06.99)

(54) Title: RIBOZYMES CAPABLE OF CONFERRING RESISTANCE TO POTYVIRUS INFECTION, AND PLANTS EXPRESSING SAID RIBOZYMES

(57) Abstract

The invention concerns a nucleic acid molecule, called polyribozyme, having endoribonuclease activity, comprising at least first and second hybridising regions and at least two catalytic regions, said first and second hybridizing regions being complementary to first and second target regions, respectively, in the genomic (+) RNA of a potyvirus, or in the replicative (–) strand RNA of a potyvirus, said first and second target regions being either: within or comprising different genes or untranslated regions on the same RNA strand, or on different RNA strands, or within a region corresponding to one gene or cistron on the replicative (–) strand RNA.

RIBOZYMES CAPABLE OF CONFERRING RESISTANCE TO POTYVIRUS
INFECTION, AND PLANTS EXPRESSING SAID RIBOZYMES.

The present invention relates to a class of « ribozymes », i.e. nucleic acid molecules having endoribonuclease activity, which are capable of conferring on plants, resistance to infection by potyviruses. The invention also relates to plants and plant cells expressing the ribozymes, and showing resistance or tolerance to potyvirus infection.

Potyviruses belong to the picorna-like supergroup of viruses, and within this supergroup, to the family of the potyviridae. Potyviruses are mostly aphid-, transmitted, although some are transmitted by mites, whiteflies or fungi. They are one of the world's most widespread plant pathogens, causing especially severe disease in cucurbits in many countries.

Potyviruses include : Zucchini yellow mosaic virus (ZYMV), Watermelon mosaic virus 2 (WMV2), Tobacco vein mottling virus (TVMV), Tobacco etch virus (TEV), Potato virus Y(PVY), Pepper mottle virus (PepMoV), Soyabean mosaic virus (SbMV), Papaya ringspot virus (PRSV), Pea seed-borne mosaic virus (PSbMV), Turnip mosaic virus (TuMV), Johnson grass mosaic virus (JGMV), Plum pox virus (PPV, Maize Dwarf Virus (MDV), Bean common mosaic virus (BCMV), Bean yellow mosaic virus (BYMV), Carnation vein mottle (CarVMV), Celery mosaic virius (CeMV), Clover yellow vein virus (ClYVV), Henbane mosaic virus (HMV), Lettuce mosaic virus (LtMV), Sugarcane mosaic (SCMV).

Potyviruses regarded as strains of the above potyviruses, or of uncertain relationship to the above viruses, and included in the definition of « Potyviruses » for the purposes of the present invention, include : Clover mosaic virus, Datura mosaic

2

virus, Desmodium mosaic virus, Freesia mosaic virus, Mungbean mosaic virus, Passionfruit ringspot virus, Pea necrosis virus, Sweet potato A virus, Wild potato mosaic virus. Possible members of the Potyvirus family include : Carrot mosaic virus, Celery latent virus, Palm mosaic virus and fern virus.

Of these viruses, WMV2, ZYMV and PRSV are responsible for severe yield reduction in Cucurbit crops. Complete resistance to infection by potyviruses, rarely occurs in economically important crops, although tolerance, in which plants are susceptible to infection but either remain symptomless or only slowly develop inconspicuous symptoms and often support restricted virus replication, is more common.

The establishment of a means to impart resistance to potyvirus infection and/or disease development would clearly have considerable economic importance.

In addition to breeding for resistance, a number of approaches using genetic engineering have been explored in an attempt to provide resistance to potyviruses.

For example, Van der Vlugt (Plant Molecular Biology, 1992, vol. 20 : 631-639) produced transgenic tobacco plants expressing both translatable and non-translatable Coat Protein (CP) RNA from PVY. Full protection (i.e completely free of systemically spread virus) following mechanical inoculation of transformed tobacco plant was observed for some plants, whether or not the CP RNA was translated. The work described did not extend to testing for resistance against inoculation by aphids, or against other related and unrelated potyviruses.

Immunity to infection by ZYMV in transgenic muskmelon plants expressing the full-length ZYMV CP gene has been reported by Fang and Grumet (Mol. Plant-Microbe Interactions, 1993, vol. 6 n°3 : 358-367).

3

Protection was also exhibited by plants expressing either a truncated, conserved core of the CP gene, or CP antisense RNA, although these constructs were reported to be less efficient then the full-length sense CP gene. These results are in contrast to those obtained by Lindbo and Dougherty (Mol. Plant-Microbe Interactions, 1992, vol. 5 : 144-153) where expression of antisense and translationally deficient sense TEV CP RNA resulted in much higher protection against TEV infection than that conferred by expression of the sense version of the full-length CP gene. The effectiveness of coat protein and antisense techniques in obtaining potyvirus resistance therefore appears to have met with varying success depending upon the precise constructs and viruses involved.

Another approach has been described by Van der Vlugt et al (Virus Research, 1993, vol. 27 : 185-200) who synthesized two ribozymes against a relatively conserved region of the RNA-dependent, RNA-polymerase-encoding cistron (also known as the NIb cistron) of PVY. The target triplet of these ribozymes is GUC at position 8122-8124 of PVYn as described in Robaglia (1989). These ribozymes were tested _in vitro_ against an artificial PVY-specific target RNA. Whilst the « long-arm » ribozyme (total arm length 106 bases) cleaved the substrate RNA efficiently, the « short-arm » ribozyme (total arm length 19 bases) failed to cleave. This short-arm ribozyme was however capable of cleaving a much shorter PVY specific RNA transcript. The authors conclude that the cleavage activity of the ribozyme is in part determined by the substrate involved, long substrate RNA's such as full-length viral RNA's being likely to possess extensive secondary structure which prevents the formation of the correct ribozyme hammerhead structure and thus prevents cleavage. No _in vivo_ tests are described for these ribozymes.

4

The technical problem underlying the present invention is to provide a means for conferring protection, on plants, to potyvirus infection, or disease development. Advantageously, the means provide protection to a large number of host plants against a particular potyvirus and to serologically related potyviruses, preferably in the heterozygous state.

The present invention provides a solution to the above technical problem by providing nucleic acid molecules or « ribozymes » having endoribonuclease activity which are designed to specifically cleave full-length potyviral RNA, (either (+) sense genomic RNA and/or the (-) sense replicative strand RNA).

The ribozymes of the present invention may comprise single or multiple catalytic regions (« mono- » or « poly- » ribozymes).

The present invention relates to a nucleic acid molecule, called « ribozyme », having endoribonuclease activity, comprising at least one hybridising region and at least one catalytic region, said hybridising region being complementary to one or more target regions of the genomic (+) RNA of a potyvirus, and/or to one or more target regions of the replicative (-) strand RNA of a potyvirus. Advantageously, the said target region does not contain the GUC triplet located at position 8122-8124 of the PVYn strain or the equivalent triplet of other potyviruses.

More particularly, the invention relates to a nucleic acid molecule, called polyribozyme, having endoribonuclease activity, comprising at least first and second hybridising regions and at least two catalytic regions, said first and second hybridizing regions being complementary to first and second target regions, respectively, in the genomic (+) RNA of a potyvirus, or in the replicative (-) strand RNA of a

5

potyvirus, said first and second target regions being either :
- within or comprising different genes or untranslated regions on the same RNA strand, or
- on different RNA strands, or
- within the same gene on the replicative (-) strand RNA.

The resistance provided by the ribozymes of the present invention is, for a number of reasons, unexpected on the basis of previous knowledge. Specifically, it has been shown both by the present inventors and by other authors (Edington and Nelson, Gene Regulation - Biology of Antisense RNA and DNA : Ed. Erickson and Izant, Raven Press Ltd. New York, 1992), that _in-vitro_ cleavage results obtained with ribozymes are not always predictive of _in-vivo_ situations. _In vitro_ data obtained by Van der Vlugt (1993 _supra_) can therefore not be reliably extrapolated to _in vivo_ situations. Furthermore, use of ribozymes on full-length viral RNA could not be expected to work efficiently in view of the secondary structure associated with such long RNA, as pointed out by Van der Vlugt et al (1993 et al _supra_).

Moreover, some authors have postulated that the (+) strand RNA may be protected by a protein coat during the early stages of infection (« stripeosome » hypothesis) until replication, and therefore would not be an appropriate target for ribozymes.

Many of the present ribozymes give rise to resistance in the heterozygous state, thereby facilitating transmission of the resistance characteristic to progeny of transgenic plants.

In addition, the polymeric ribozymes of the invention are capable of conferring resistance _in vivo_ to virus strains related to, but different from the particular targeted strains.

6

Finally, many of the polymeric ribozymes of the invention target sites which are either on different strands of RNA (i.e (+) and (-)), or have sites which are widely separated regions on the same strand. The ribozymes thus do not only act as a single molecule simultaneously targeting a plurality of sites, but rather can act as if the molecule were a plurality of monoribozymes each targeting a single site, thus increasing the probability that the substrate will be cleaved.

In the context of the present invention, the following terms have the following meanings :

Potyvirus : A large group of PLANT VIRUSES in which the virions are flexuous filaments (ca. 680-900 x 11 mm) comprising one type of coat protein (MWt ca. 32,000-36,000) and one molecule of linear positive-sense ssRNA (MWt ca. 3.0-3.5 x 10^6) comprising a single large open reading frame (ORF) coding for a polyprotein of ca. 350.000 Da which is proteolytically processed by virally encoded proteinases.

As a group, the viruses infect a wide range of host plants, including monocots and dicots, but individual members typically have a narrow host range. Symptoms depend on plant cultivar, virus strain, and environmental conditions, but typically include mosaic or mottle in the leaves, colour-breaking in flowers, mottling and/or distortion in the fruits ; characteristic proteinaceous inclusion bodies (serologically unrelated to virus coat protein) are formed in the cytoplasm of infected plant cells, appearing as « pinwheels » in transverse section or as « bundles » in longitudinal section. A few members also form crystalline nuclear inclusions. Potyviruses are transmitted (non-persistently) by aphids and can be transmitted mechanically ; some viruses tentatively

included in the group are transmitted by whitefly,
mites or fungi.

The replication cycle of potyviruses involves the
following successive steps : infection, uncoating,
translation, polyprotein processing to yield
functional, viral coded products, genome replication
via a negative strand RNA template, progeny virus
particle assembly, cell-to-cell movement within the
infected plant and vector transmission from plant to
plant.

Gene : In the context of potyviral genomic RNA, the
term « gene » as used herein is synomynous with
« cistron ». All potyviral proteins are generated from
a single large primary translation product
(polyprotein) by viral proteases, produced by
translation of the single genomic ORF. In the present
context, « gene » signifies a region of the genomic ORF
coding for a protein released as an independent
molecule by viral proteases. Potyviral « genes » or
cistrons are referred to herein by reference to the
designation accepted in the art for the type members of
the potyviruses family PVY as shown in figure 20, or
their functional equivalent in other members of the
family. Reference to genes or cistrons in the (-)
strand signifies the region corresponding to (i.e.
complementary to) that gene or cistronin the (+)
strand.

(+) sense RNA : In the context of potyviral RNA, « (+)
sense RNA » signifies the genomic RNA of the virus as
contained within the virion, which serves directly as a
template for translation of a polyprotein. The (+)
sense RNA may be encapsidated or not. Naked (+) sense
RNA is also infectious.

8

<u>Replicative or (-) sense RNA</u>: In the context of potyviral RNA, « (-) sense or replicative RNA » signifies the full-length, non-coding negative strand template RNA produced by the copying of the entire (+) RNA genome by a viral polymerase (NIb). No sub-genomic RNAs are produced in potyviruses. In the context of the present invention, positions of nucleotides in the viral (-) sense RNA are indicated by reference to the corresponding position of the (+) strand, with a (-) in brackets. Reference to genes or cistrons in the (-) strand (which is in fact non-coding) signifies a region or a site in the (-) strand corresponding to that region or site in the (+) strand.

<u>Potyviral RNA</u> : In the present context, this expression encompasses both (+) sense genomic RNA and (-) sense replicative RNA.

<u>Resistance</u> :reduction or prevention of viral infection e.g. absence of symptoms and absence of detectable viral particles, absence or reduction or delay of apparition of symptoms. Symptoms are sometimes atypical. Recovery possible.

<u>Hybridising /complementary</u> : the « arms » of the ribozyme are said to « hybridise » to a target RNA sequence or to be « complementary » to the target region in the potyvirus. Such hybridisation and complementarity is defined herein as an extent of duplex formation sufficient to form a stable enough structure so that cleavage adjacent to the target triplet can occur. Depending on the target sequence, the ribozyme and the secondary structure of the RNA, preferably less than about 5-10, more preferably less than about 3 to 5 and even more preferably less than 1 to 3 bases would be mismatched between the target RNA

9

sequence and one or both hybridising arms of the ribozyme. Over the whole length of the hybridising arm, not more than 10% of mismatches should occur. Accordingly, the present invention extends to ribozymes described herein, to the polyribozymes containing them, and to any or all functional mutants, derivatives, parts, homologues or analogues thereof. Examples of such mutants include single or multiple nucleotide substitutions, deletions and/or additions to the hybridising regions and/or catalytic regions of the ribozymes provided such modified ribozymes are still functional in cleaving the target RNA sequence, and/or conferring resistance to potyviruses. Such mutants are said to exhibit « substantial complementarity » to the viral target sequence.

The present invention therefore extends to a ribozyme having one or a plurality of hybridising domains and catalytic domains, wherein the hybridising domain comprises a sequence of nucleotides substantially complementary to a sequence of nucleotides in a potyvirus RNA in vitro under high stringency conditions and wherein said catalytic region is capable of cleaving said target RNA sequence. This aspect of the present invention is conveniently determined by the extent of cleavage of the target RNA sequence in vitro.

For the purposes of the present invention, the stringency conditions set forth in Sambrook et al, Molecular Cloning : A laboratory Manual, Cold Spring Harbor laboratories, Cold Spring Harbor, NY, USA, pages 387-389 at paragraph 11 are considered high stringent conditions. Alternatively, medium stringency may be used involving 0.25-0.5 % w/v SDS at 45°C for 2-3 hours. The alternative conditions are applicable depending on concentration, purity and source of nucleic acid molecules.

Hybridising region : the total length of hybridizing
arms. Normally, in the context of the invention, when a
hybridizing region is said to be complementary to a
region within or comprising a gene, this is intended to
signify that the site at which endoribonucleolytic
cleavage occurs is within that gene.

Target region : In the context of potyviral RNA,
« target region » signifies a stretch of
ribonucleotides, normally having a length of from
approximately 10 bases up to the length of the full
potyviral RNA, to which the hybridising region of the
ribozymes of the invention hybridise. The target region
contains a precise target site at which cleavage can
occur when acted upon by the ribozyme. Generally
speaking, the extent of the « target region » is
defined as the length of the sequence in the viral RNA
over which hybridisation occurs.

Target triplet : In the context of hammerhead
ribozymes, « target triplet » signifies the three bases
immediately 5' to the cleavage site in the substrate
RNA. The target triplet will be noted herein as X·X·X·
wherein X· = any ribonucleotide, preferably X·UX·
wherein X· is any ribonucleotide and U is uracil, or
X·UY wherein Y is adenine, cytosine or uracil.

Target sites in(-) sense RNA : Ribozyme target sites
in the potyviral (-) strand replicative RNA are
designated by reference to the corresponding cistron
(or « gene ») in the (+) strand.

     The potyviruses targeted by the ribozymes of the
present invention are preferably :
     Zucchini yellow mosaic virus   (ZYMV),

Watermelon mosaic virus 2	(WMV2),
Tobacco vein mottling virus	(TVMV),
Tobacco etch virus	(TEV),
Potato virus Y	(PVY),
Pepper mottle virus	(PepMoV),
Soyabean mosaic virus	(SbMV),
Papaya ringspot virus	(PRSV),
Pea seed-borne mosaic virus	(PSbMV),
Turnip mosaic virus	(TuMV),
Johnson grass mosaic virus	(JGMV),
Plum pox virus	(PPV),
Maize Dwarf Virus	(MDV).

Strains of ZYMV and WMV2 as well as serologically related viruses are particularly preferred.

Target regions of potyviral RNA which are particularly preferred are those cistrons (and the corresponding regions of (-) sense RNA) which give rise to proteins involved in viral processing, replication, and assembly such as NIa, NIb and CP. These cistrons are located in the 3'-half of the genomic RNA.

According to a preferred embodiment, the nucleic acid molecules of the invention comprise at least one unit « R » having the general formula :

$$3' \; [(X)_n - Q - (X')_{n'}]_p \; 5'$$

wherein :

- X and X' each represent, independently, any ribonucleotide ; $(X)_n$ and $(X')_{n'}$ represent ribonucleotide sequences substantially complementary to at least part of said target region of potyvirus genomic (+) RNA or to at least part of its (-) strand replicative RNA ;

- n and n' are each equal to or greater than 4 ;

- Q represents the catalytic region of a ribozyme ;

- $p \geq 1$.

12

In the above general formula $(X)_n$ and $(X')_{n'}$ represent the complementary sequences or « arms » of the ribozyme which hybridise to the chosen target region of the potyvirus. $(X)_n$ and $(X')_{n'}$ together represent the « hybridizing region » of the ribozyme. $(X)_n$ and $(X')_{n'}$ normally hybridise to the sequences immediately 5' and 3' respectively of the target site in the viral RNA.

The values of n and n' are, independently, at least 4, preferably at least 8, for example from 4 to 4000. Preferred values are from 4 to 50, for example 6 to 50 or 8 to 12 nucleotides for « short arm » ribozymes. The total length of the two hybridising arms is at least 8 nucleotides, preferably at least 16. For « long arm » ribozymes, n and n' normally have values greater than 50, for example 50 to 400, particularly 60 to 300 nucleotides.

In the above general formula, p = 1 for monoribozymes. For polyribozymes, $p \geq 2$, for example 2-20 or 4-10. Polyribozymes are particularly preferred.

Preferably, the complementary sequences $(X)_n$ and $(X')_{n'}$ are complementary to target regions of potyvirus RNA which are conserved between strains of a given species, and preferably also between different species of potyvirus. In the context of the present invention, « conserved » signifies that the sequences on at least one side, and preferably on both sides of the cleavage site in the substrate are essentially identical over a length of 20 bases, i.e. have at least 85% identity, and preferably at least 90% identity over 20 bases. Generally speaking, in a conserved target region, the target triplet itself is identically conserved between two viruses. Exceptionally, since the third base in the target triplet X·X·X· remains unpaired after hybridisation of the ribozyme to the target region in the substrate, this base may vary between two viruses

in a conserved region without affecting cleavage capacity.

According to yet a further embodiment of the invention, the hybridising arms of the polyribozyme may be derived from at least two different types of virus. Such polyribozymes are referred to hereinafter as « chimeric polyribozymes ». As an example of this type of construction, mixtures of ZYMV and WMV2 derived sequences can be cited. This type of polyribozyme ensures the inactivation of both types of virus, if present together, or allows the use of a single construction in different geographical areas.

According to a preferred embodiment, the catalytic region « Q » in the above general formula is a hammerhead ribozyme catalytic region. Such ribozymes are described in European patent EP-B-0 321 201. Such a ribozyme can be represented as a molecule of formula (I):
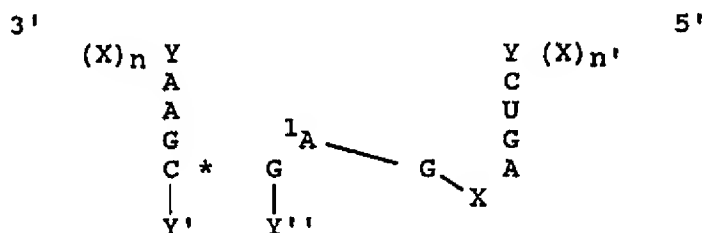
```
   3'                                          5'
     (X)n  Y                            Y  (X)n'
           A                            C
           A                            U
           G           1A               G
           C  *    G  ─────  G          A
           |        |            X
           Y'       Y''
```

wherein, $(X)_n$ and $(X)_{n'}$ represent ribonucleotide sequences substantially complementary to an RNA sequence of a potyvirus (either genomic or replicative intermediate) ; an (*) represents a base pair between complementary ribonucleotides ; Y represents any ribonucleotide and each Y may be the same or different ; y' and Y" represent ribonucleotides of complementary sequence along at least part of their length to allow base pairing between the oligoribonucleotides, or Y' and Y" together form a single RNA sequence wherein at least part of said

sequence comprises a stem formed by base pairing between complementary nucleotides ; and optionally, an additional nucleotide selected from any one of A, G, C or U is inserted after ^{1}A.

In another embodiment, the ribozyme is as set forth in formula (II) :
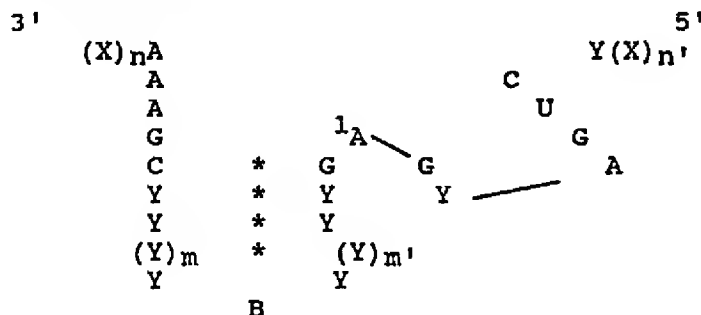
```
   3'                                        5'
  (X)nA                               Y(X)n'
     A                            C
     A                          U
     G                  1A          G
     C      *   G      G          A
     Y      *   Y         Y ____
     Y      *   Y
  (Y)m      *  (Y)m'
     Y              Y
           B
```

wherein $(X)_n$ and $(X)_{n'}$ represent ribonucleotide sequences substantially complementary to an RNA of a potyvirus ; Y represents any ribonucleotide and each Y residue may be the same or different ; an (*) represents a base pair between complementary ribonucleotides ; n and n' are as previously defined ; B represents a bond, a base pair, a ribonucleotide, or an oligoribonucleotide containing at least 2 ribonucleotides ; m and m' are 1 or more and may be the same or different ; and optionally, an additional nucleotide selected from any one of A, G, C or U is inserted after ^{1}A.

In a preferred embodiment, the ribozyme is as set forth in formula (III) :

```
   3'                                  5'
  (X)nA                        Y(X)n'
     A                        C
     A                        U
     G              A ____      G
     C     *   G          G   A
     A     *   U         U /
     G     *   C
     G     *   C
     A ____        G
     G ____   U
```

15
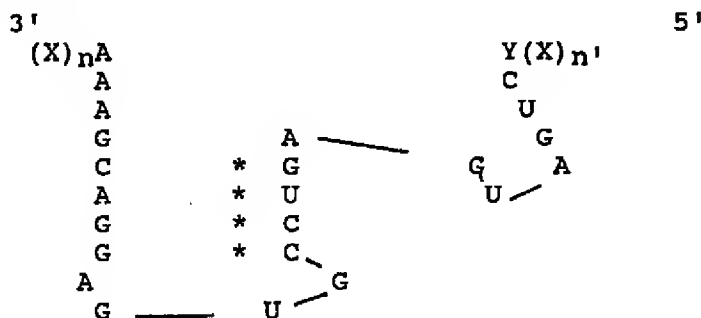
wherein $(X)_n$ and $(X)_{n'}$ represent a ribonucleotide sequence substantially complementary to an RNA of a potyvirus ; Y represents any ribonucleotide and each Y residue may be the same or different ; an (*) represents a base pair between complementary ribonucleotides ; n and n' are as previously defined.

According to further embodiments of the invention, the catalytic region of the ribozyme, for example "Q" in the above general formula may be a hairpin domain as described in WO 95/06731 and EP-B-360257. An example of such a hairpin domain is given below in Formula IV :
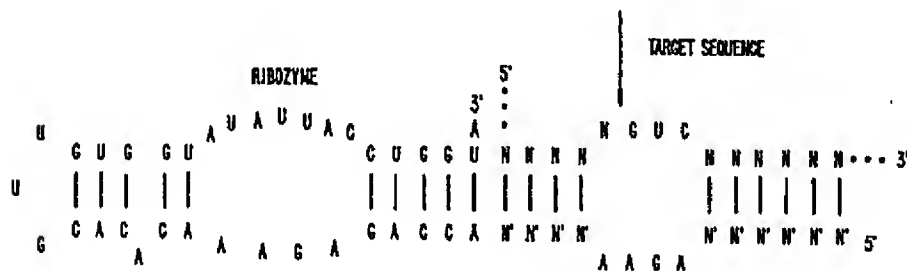
Alternatively, the catalytic region of the ribozyme may be the catalytic domain of a Hepatitis delta virus, as described in WO 95/06731, or the catalytic domain of an RnaseP ribozyme as described in EP 546092.

Although the catalytic regions illustrated above have a conserved structure and sequence, it has been observed that some nucleotides may be deleted, inserted, substituted or modified without prejudice to the activity of the ribozyme. The invention comprises the use of these modified catalytic regions in the polyribozyme provided that their catalytic activity is conserved. This activity can be verified by using the tests described below.

16

For example, one or more nucleotides of the catalytic region II illustrated in formula I, II and III may be replaced by nucleotides containing bases such as adenine, guanine, cytosine, methylcytosine, uracil, thymine, xanthine, hypoxanthine, inosine or other methylated bases. The « conserved » bases C-G which together form the first base pair of the catalytic loop, can be replaced by U-A.

Other preferred variant hammerhead catalytic regions have the following sequences :

Mutant ribozyme I :　　　　　　cugaugaGUCCugaGGACgaa
Mutant ribozyme II :　　　　　　cugaugaGUCCcgugaGGACgaa
Mutant ribozyme III :　　　　　cugaugaGUCgugagGACgaa
Mutant ribozyme IV :　　　　　　cugaugaGCGgugaCGCgaa.

Wild type sTobRV :　　　　　　　cugaugaGUCCgugaGGACgaa
(Helix II shown in upper case letters)

Because ribozymes act in a sequence-specific manner,it is possible for a target RNA to become resistant to ribozyme-mediated inactivation by just changing one or a few critical bases. For the present application, it was also desirable to build resistance genes which would be protect plants against infection by several strains of a virus species. Both the goals of achieving a robust resistance gene , and obtaining protection against many strains, were achieved by selecting the target sites for the ribozymes according to the following criteria :
- sites are chosen which cannot easily mutate without changing the coding capacity (i.e. the amino acids corresponding to a particular codon) ;
- sites are chosen which are located in highly conserved regions.

17

In a preferred embodiment of the invention, the 3' $(X)_n$ sequence of a given R unit and the 5' $(X')_{n'}$ sequence of the adjacent R unit are complementary to sequences which are contiguous in the potyvirus genomic (+) strand RNA or replicative (-) strand RNA.

According to this variant, the complementary sequence may hybridise with the entire length of the target transcript. On the other hand, it may hybridise with only a portion or region of the targeted transcript. The portion or region in question must be long enough to allow the inclusion of at least 1 catalytic region in the corresponding sequence of the polyribozyme. In general, for this variant, the length of the complementary sequence, not counting the catalytic regions (the sum of the hybridising arms), may vary from about 4 to 4000 bases. Lengths of 40 to 1000 are particularly preferred.

According to another highly preferred embodiment of the invention, in the polyribozyme having the formula :

$$R_1-R_2-R_3-------R_{n'}$$
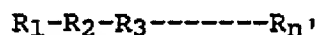
the 3' hybridising arm $(X')_n$ of a given « R » unit and the 5' hybridising arm $(X')_{n'}$ of the adjacent « R » unit are complementary to non-contiguous portions of the genome or of the (-)sense RNA transcript. An example of this type of structure is the « short arm polyribozyme ».

The structure of these non-contiguous polyribozymes can be thought of as a sequence complementary to the target sequence in which ribozyme catalytic regions are embedded and which has undergone deletions in the complementary sequence. According to this embodiment, the distance between two catalytic regions in the polyribozyme is normally shorter than

18

the distance between the two corresponding X·X·X· sites
in the target sequence. The distance between ribozymes
can be reduced significantly with respect to the
distance between the corresponding target sites in the
substrate, for example, if two adjacent targeted X·UX·
sites in the substrate are separated by a distance of
1000 nucleotides, the distance between the catalytic
regions of the « short arm ribozyme » may be as little
as 30 or 40 nucleotides. Reductions of over 90% of the
length of the hybridising sequence may therefore be
made. The minimum length between two adjacent catalytic
regions is normally about 8 nucleotides and preferably
at least 20. The present inventors have shown that such
shortened hybridising sequences do not, surprisingly,
prevent the polyribozyme from efficiently cleaving the
substrate and can, in fact, in many cases improve the
efficiency of the inactivation.

For example, the reduction in the length of the
complementary sequence avoids problems associated with
the formation of competing or unfavourable secondary
structure leading to the inactivation of the ribozyme.
Thus, for a given length of ribozyme molecule, more
catalytic regions may be inserted targeting more
cleavage sites in the substrate.

Furthermore, the use of the short arm
polyribozymes means that efficent polyribozymes may be
constructed even if only parts of the sequence of the
substrate are known. It is sufficient to know the 10 or
15 bases surrounding an X·X·X· target site.

Another important advantage is the possibility to
use, as the shortened hybridising arms, exclusively
those sequences which are conserved in different
viruses, for example viruses of the same type. In this
way, different viruses may be inactivated by the same
polyribozyme even if the sequences between the
conserved regions are totally dissimilar. These viruses

19

may be of the potyviridae family or species within the family or different isolates of the same species.

Finally, the smaller molecule allows an easier and biologically more economical synthesis. The functional nature of the « short arm » ribozymes of the invention proves that, in a polyribozyme, the different catalytic regions can act independently of each other.

Long-arm ribozymes, having arm lengths of at least 50 nucleotides, either contiguous or non-contiguous, are also preferred embodiments of the invention. Examples are shown in Table 8 below.

Any one polyribozyme of the invention may include a mixture of « long arm » and « short arm » ribozymes, and also of « contiguous » and « non-contiguous » ribozymes. Alternatively, it may include only long arm ribozymes, or only short arm, or only contiguous, etc... The structure of the ribozyme can therefore be optimised by testing using the in vitro - in vivo and transient assay described below.

According to a further variant of the invention, the polyribozymes may include, between two consecutive « R » units, a non-complementary sequence (A) which has a length of at least 1 base, for example 3 or more bases, and which is non-complementary to the (+) or (-)sense potyviral RNA.

The non-complementary sequence of the polyribozyme may have a precise function, for example, it may be constituted by a coding sequence which can be used to select transformants or also a sequence containing a ribozyme which acts on a substrate other than the primary target of the DNA virus or which is cis-acting on a part of the polyribozyme. It typically contains a polylinker for cloning, or an adaptor to facilitate constructions of chimeric polyribozyme. The non-complementary sequence « A » usually has a length comprised between about 2 and 500 bases, for example 20

to 100 bases. When there is a plurality of non-complementary sequences, they may together constitute as much as about 90% of the length of the polyribozyme, more preferably 50%.

Preferred ribozymes of the present invention have multiple catalytic domains (« polyribozyme ») targeting different regions or different RNA strands of the potyvirus. According to their construction (long arm or short arm, contiguous or not contiguous), the polyribozyme may act as a series of individual monoribozymes or may act as a « unimolecule » simultaneously against each of its target sites.

In particular, according to a preferred variant, in a polyribozyme, two ribozyme units R1 and R2 target either, different cistrons (or « genes ») on the same RNA strand, or they target the same or different cistrons on different RNA strands. In such systems, the targeted RNAs differ from each other, either in that they comprise, or are within different cistrons, and/or in that they are within different strands. Alternatively, the targeted sequences may be within the same cistron on the replicative (-) strand RNA. Specifically, such a ribozyme comprises at least two of said R units, « R_1 » and « R_2 », wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of R_1 are complementary to sections of a first target region, Z_1, of the potyvirus genomic (+) RNA or of the replicative (-) strand RNA, region Z_1 being within or comprising a first cistron or untranslated region (UTR), and X and X' of R_2 are complementary to sections of a second target region, Z_2, of the potyvirus genomic RNA or (-) strand replicative intermediate, region Z_2 being within or comprising a second cistron or untranslated region, Z_1 and Z_2 being either :
- different cistrons or untranslated regions, on the same RNA strand, or

21

- being the same or different cistrons or untranslated
regions on different strands of the potyvirus RNA, or
- being the same genes on the (-) strand replicative
RNA.

Regions Z_1 and Z_2 contain the target triplet for
the ribozymes R_1 and R_2 respectively. Regions Z_1 and
Z_2 thus contain the ribozyme cleavage sites of R_1 and
R_2 respectively.

Preferred examples of this type of polyribozyme
are those comprising a plurality of targets on the (+)
strand, but within different cistrons. This type of
polyribozyme is exemplified by a nucleic acid wherein
the complementary sequences $(X)_n$ and $(X')_{n'}$ of R1 are
complementary to a region within or comprising a first
gene or U.T.R. on the genomic (+) RNA, and X and X' of
R_2 are complementary to a region within, or comprising,
a second gene or U.T.R. on the genomic (+) RNA.

Alternatively, all targets may be on the (-)
strand, usually within regions corresponding to
different cistrons or untranslated regions. This type
of polyribozyme is exemplified by a nucleic acid
wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of
R_1 are complementary to a region within or comprising a
region corresponding to a first gene or U.T.R. on the
replicative (-) strand RNA, and X and X' of R_2 are
complementary to a region within or comprising a region
corresponding to a second gene or U.T.R. on the
replicative (-) strand RNA.

Another example of the ribozyme of the invention
is a polyribozyme which targets regions on different
RNA strands of the potyvirus. This type of polyribozyme
is exemplified by a nucleic acid wherein the
complementary sequences $(X)_n$ and $(X')_{n'}$ of R_1 are
complementary to a region Z_1 within or comprising a
gene or U.T.R. on the genomic (+) RNA, and X and X' of
R_2 are complementary to a region Z_2 within or

22

comprising a region corresponding to a gene or U.T.R. on the replicative (-) strand RNA. In this type of polyribozyme, the target regions may correspond to (+) and (-) forms of the same cistron, or they may be regions of the (+) and (-) strand which have no functional or structural relation to each other. When the target regions correspond to (+) and (-) forms of the same cistron, the polyribozyme may be defined as a nucleic acid of the invention wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of R_1 are complementary to a region Z_1 within or comprising a cistron « C » on the (+) strand RNA and the complementary sequences of $(X)_n$ and $(X')_{n'}$ of R_2 are complementary to a region Z_2 on the (-) strand within or comprising the region corresponding to the same cistron « C ».

In all of the above ribozymes, the complementary sequences $(X)_n$ and $(X')_{n'}$ may be long or short arms as previously defined. The polyribozymes may comprise both long and short arms. Polyribozymes targeting different cistrons on the same strand are frequently designed to be long arm structures, except when the targets are all the (-) capsid protein gene in which case short arm structures are preferred. Polyribozymes targeting the same cistrons on different strands can be short arm, or a mixture of short and long arm, or alternatively long arm. Preferred polyribozymes have at least one target site in the capsid protein gene (+) or (-) strand, in association with at least one target site in a different gene.

A further preferred polyribozyme of the invention is a polyribozyme targeting at least two sites within the same cistron on the same strand. This type of polyribozyme comprises at least two R units, « R_1 » and « R_2 », wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of R_1 are complementary to sections of a target region Z_1 within a given gene or untranslated region of

23

the potyvirus genomic (+) strand RNA or of the corresponding region of the replicative (-) strand RNA, and the hybridising sequences (X) and (X') of R_2 are complementary to a target region Z_2 within the same gene or untranslated region as that of Z_1, on the same RNA strand. Preferably, both regions Z_1 and Z_2 are on the (-) strand. However, if, according to this variant, Z_1 and Z_2 are both on the genomic (+) strand and both are within the capsid protein gene, the polyriboyme contains a further catalytic unit targeting a site on a different gene or on the (-) strand RNA.

Preferred target regions for the polyribozymes of the invention are regions within or comprising the cistron giving rise to the capsid protein, the NIa cistron, the NIb cistron, the 3' untranslated region of a potyvirus, or the corresponding sequences in the replicative (-) strand RNA.

Particularly preferred polyribozymes include amongst their target sites, at least one within the capsid protein cistron, or within the corresponding cistron of the (-) strand RNA. Such ribozymes include nucleic acid molecules wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of at least one R unit are complementary to regions within or comprising the capsid protein (CP) gene, or the corresponding sequences in the replicative (-) strand RNA. Such polyribozymes may also include ribozyme sequences targeting other cistrons or non-translated regions of a potyvirus RNA or its replicative (-) strand RNA. Polyribozymes targeting at least one site in the region of the (-) strand RNA corresponding to the capsid protein are especially preferred, particularly in association with a target site in a different gene. The (+) or (-) strand polyribozymes targeting the capsid protein may have any combination of other targets, arm lengths and arm configurations described above.

24

Preferred examples of capsid protein ribozymes have either just one target in the capsid protein in the (+) or (-) strand, or alternatively three or four targets in the capsid protein (+) strand RNA. Capsid protein ribozymes according to this variant of the invention are especially advantageous when complementary to WMV2.

Specific examples of ribozymes of the invention are the following :

- ribozymes comprising at least two R units, R_1 and R_2, wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of R_1 and of R_2 are respectively complementary to regions within the CP cistron (either (+) or (-) strand) and the 3' non-translated region of the genomic (+) strand RNA ;

- ribozymes comprising at least two R units, R_1 and R_2, wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of R_1 and R_2 are respectively complementary to regions within the region of the replicative (-) strand RNA corresponding to the CP cistron and the NIb cistron, or to the NIb cistron and the NIa cistron ;

- ribozymes comprising at least two R units, R_1 and R_2, wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of R_1 are complementary to the CP cistron of the genomic (+) strand RNA, and the hybridising regions of R_2 are complementary to all or a portion of the region corresponding to the CP cistron on the replicative (-) strand RNA ;

- ribozymes comprising at least three R units, R_1, R_2 and R_3, wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of R_1 are complementary to regions within the CP cistron or the 3'NT region of the genomic (+) strand RNA, the complementary sequences $(X)_n$ and $(X')_{n'}$ of R_2 are complementary to regions within the region corresponding to the NIb cistron on the replicative (-) strand RNA and the complementary sequences $(X)_n$ and $(X')_{n'}$ of R_3 are complementary to regions within the

region corresponding to the NIa or CP cistrons on the
replicative (-) strand RNA ;
-ribozymes wherein the complementary sequences $(X)_n$ and
$(X')_{n'}$ of R_1 and R_2 are complementary to regions within
the CP gene on the replicative (-) strand RNA.

Further examples of preferred ribozymes are
indicated in table 8 below.

In addition to the choice of target site(s),
target region(s) and catalytic region(s), a further
variant can be obtained by incuding in the ribozyme
construction, all or part of a 3'non-coding region of a
potyvirus, preferably WMV2, to control intracellular
localisation of the ribozyme gene. This may help to
direct the ribozyme RNA to the cellular compartment
where the potyviral RNA might be. The enhanced
functioning of ribozymes carrying the 3' non-coding
region is suspected by analogy to the observations in
animal cells (Sullenger and Cech 1993, Science vol. 262
p 1566-1569) that a ribozyme against a retrovirus was
much more efficient when it was co-localised with the
target, using a retroviral vector to express it.

The 3' non-translated sequence usually contains
all the untranslated region of the virus and may also
contain the last 10 to 20 codons of the protein and the
first 10 to 15 residues of the viral poly A tail.

The ribozymes of the present invention can be
tested for their capacity to cleave a substrate RNA,
and / or to confer resistance to potyvirus infection,
by <u>in vitro</u>, <u>in vivo</u> or transient assays as described
below in examples 8, 9 and 10 respectively.

The in vitro test involves incubating a suitable
RNA substrate, uniformly labeled with a radioactive
label, with the ribozyme under test, prepared by in
vitro transcription. Incubation is carried out at
varying temeratures as described in example 8. The
capacity of the ribozyme to cleave the substrate at the

26

expected target sites is shown by carrying out an autoradiogram on a polyacrylamide gel.

The ribozymes and genetic constructs of the present invention can be conveniently tested _in vivo_ using virus-sensitive plants. The analysis may be qualitative such as by screening for severity of symptoms, or quantitative where samples are analysed, for example, by measuring the level of virus replicating in infected plants using procedures such as ELISA, or alternatively, the level of ribozyme RNA, or the level of target RNA. A combination of qualitative or quantitative analyses may also be employed.

The genetic nature of the virus-resistant transgenic plants can be determined by carrying out a segregation analysis. The goal of this test is to determine genetic linkage of the ribozyme gene and virus resistance. Self-fertilisation, or a cross with a non transformed genotype can be carried out on a primary descendant to obtain T_1. Subsequently, T_1 plants are inoculated with the potyvirus by mechanical inoculaion or natural inoculation or by using aphids fed on infected plants. Plants are scored for resistance to infection. A segregation ratio of 3:1 (resistant : susceptible) indicates a single locus acting as a heterozygous trait (« dominant ») ; a segregation ratio of 1:3 indicates a single locus acting as a heterozygous trait (« recessive »), and a segregation ratio of >3:1 indicates a plurality of loci, acting independently of each other, and conferring resistance as a single copy as a heterozygous trait.

The ribozymes of the invention can also be tested using a transient transformation assay as desrcibed below in Example 10. The target is linked to a reporter gene such as LUC. DNA encoding the substrate-reporet gene hybrid is co-introduced into a plant cell along

27

with a plasmid encoding the ribozyme under test. The level of expression of the reporter gene is measured. Expression is decreased when the ribozyme interacts with the target.

The present invention also extends to genetic constructs comprising the ribozymes herein described. These genetic constructs generally comprise a vector sequence as a vehicle for transferring the ribozyme to a host cell and a promoter sequence to direct transcription of the ribozyme and a terminator to signal termination of transcription. A preferred promoter includes but is not limited to a 35S Cauliflower Mosaic Virus (CaMV) or to a double 35S CaMV promoter. The CaMV promoter is described by Odell et al, Nature 313 : 810-812, 1985. The ribozymes are generally in the complementary form such that upon transcription, the ribozyme is produced. The genetic constructs of the present invention are used to generate transgenic plants capable of constitutively, developmentally or inducibly (e.g. in response to certain stimuli) expressing the ribozymes herein described. Most preferred transgenic cells and plants are rendered substantially resistant to potyvirus infection.

The invention also extends to a method for producing transgenic plants which exhibit at least some resistance to infection by potyviruses, and to the plants which can be obtained thereby. According to this aspect of the present invention, there is provided a method of generating trangenic plants exhibiting reduced symptoms following exposure to potyviruses compared to non-transgenic plants, said method comprising transforming plant cells with a genetic construct encoding a ribozyme as hereinbefore described capable of cleaving potyvirus RNA, regenerating plants from said transformed cells, self-fertilising said

28

regenerated plants and obtaining seeds there from and then growing transgenic plants from said seeds.

All of the known means for introducing foreign DNA into plants may be used, for example Agrobacterium, electroporation, protoplast fusion, bombardment with a particle gun, or penetration of DNA into cells such as pollen, the microspore, the seed and the immature embryo, viral vectors such as the Geminiviruses or the satellite viruses. Agrobacterium tumefaciens and rhizogenes constitute the preferred means. In this case, the sequence coding for the polyribozyme is introduced into a suitable vector together with all of the regulatory sequences necessary such as promoters, terminators, etc... as well as any sequence necessary for selecting the transformants.

The invention also extends to progeny, reproductive material and seeds from the transgenic plants. The ribozymes may be expressed constitutively, developmentally or inducibly in the whole plant, or in specific tissues of the plant.

As examples of plants which can be rendered resistant to potyvirus infection by the present invention, mention can be made of plants belonging to the families : Cucurbitaceae (e.g. courgette melon), Chenopodiaceae (e.g. beet), Cruciferaceae, Graminaceae (e.g. oat, wheat, sugar cane, barley, ryegrass), Leguminoseae (e.g. bean, pea, soya, peanut), Ombelliferaceae (e.g. carrot, celery, parsnip), Rosaceae (e.g. plum), Solanaceae (e.g. potato, tomato, tobacco, pepper).

The present invention also extends to the use of the claimed ribozymes in plant cells or plants, in association with other virus resistance genes or systems, either to potyviruses or to other viruses, for example, anti-CMV ribozymes (CMV signifies « cucumber mosaic virus), or coat-protein mediated systems for

29

virus resistance such as coat protein antisense or sense (translated, or translation deficient) RNA. According to this variant of the invention, the plant may comprise cells which are transformed both by the ribozyme sequences of the invention and by the sequences providing associated protection. Expression of the different sequences providing protection may be simultaneous or otherwise.

According to a further embodiment of the invention, there is provided a nucleic acid molecule which consists of a fragment of the WMV2 viral replicative (-)RNA, which has been shown to be particularly amenable to cleavage or inactivation by ribozymes, and can thus be used as an artificial target molecule. Indeed, the unexpected discovery that the sequence from the negative strand of the virus around site 1747 is specially susceptible to gene inactivation opens useful possibilities to use this sequence as a ribozyme or antisense target not only for protecting plants against viruses but for other uses as well.

This molecule, and its derivatives will be referred to hereinafter as the « artificial target » molecule.

The artificial target molecules of the invention consist of either nucleotides 1704 to 1873 of the (-) strand of the sequence illustrated in figure 18, or of at least 20, for example between 40 and 170, consecutive nucleotides of the 1704-1873 fragment, including the ATG shown at position 1748-1750 in Figure 18. The artificial target molecules may also consist of a sequence having at least 75% identity and preferably at least 85% or even 95% identity with the underlined sequence of Figure 18, or a fragment thereof hving at least 20 nucleotides.

The artififcial target sequence present on pBPL 0186 as described in figure 15, is made of

30

approximately 180 bases of viral sequence from the
minus strand of the viral genome, in the region coding
on the plus strand for the replicase NIb gene. The
sequence is made of 125 bases in 5' of the A of the GUA
ribozyme cleavage site and 44 bases in 3' of the A. The
A is the complementary base to the T shown as base n°
1748 on figure 18. Part of this sequence is present on
many RNA viruses. For example around this cleavage site
in WMV2 :

Potato Virus Y has 20 out of 22 bases identical
to WMV2 (90.9%)

Potato Virus A has 22 out of 23 bases identical
to WMV2 (95.6%)

Pea Seedborne Mosaic Virus has 20 out of 23 bases
identical to WMV2 (86.95%)

Human Cytomegalovirus has 17 out of 23 bases
identical to WMV2 (73.9%)

Influenza Virus has 17 out of 19 bases identical
to WMV2 (89.5%).

This region is therefore a good target for
molecules designed against those viruses.

The invention also relates to larger nucleic acid
molecules containing the artificial target sequences of
the invention with the exception of known full length
viral (-)RNA which naturally contains the sequence.

Preferred examples of molecules comprising the
artificial target sequences are molecules in which the
artificial target sequence is transcriptionally or
translationally fused to a second nucleic acid
molecule.

The artificial target sequence and its
derivatives is thus appended to a gene of interest
(hereafter called « the useful gene » or « gene of
interest »), in 5' or in 3' of the open reading frame
or coding sequence of the useful gene, so that it does
not prevent the function of the useful gene and makes

31

this gene easily controllable by a ribozyme (such as those described earlier), or antisense, directed against the artificial target.

The term « coding sequence » in this context includes those sequences active as RNA like tRNA, rRNA, etc. in addition to those sequences coding for proteins. The appended sequence may also be a regulatory sequences, sequences encoding functional RNA such as rRNA, tRNA, mRNA etc.

The 5' fusion can be transcriptional, or translational. Similarly the 3' fusion can be transcriptional or translational. The precise nature of the fusion depends on the functionality of the useful gene. For example, in the case of a gene coding for a protein, a translational fusion can be used if the aminoacids added to the protein (in NH_2 or COOH terminus as the case may be) do not prevent its function.

The artificial target sequence can be smaller than the sequence described above, as demonstrated by the inhibitory activity of short arm ribozymes pairing with only a very small stretch of that sequence. For example the artificial target sequence can be reduced to 40 bases on each side of the A of the GUA, or to 10 bases in 5' and 12 bases in 3' of the A, or even to 8 bases on each side of the A, if a slightly lower efficiency of the ribozyme is acceptable.

Optimal length can be found by routine experiment.

The high activity of ribozymes against the artificial target sequence in the experiments described can probably be explained in part by its accessibility to inhibitor RNA molecules. Therefore, it is possible to use not only ribozymes but also antisense RNA against the artificial target, albeit with a lower efficiency.

32

The fusion between the gene of interest and the artificial target can now be used as a target for ribozymes or antisense RNA. The inactivation of the useful gene containing the artificial target can be caused by the ribozyme gene brought in through a cross, or by the tissue specific expression of the ribozyme gene if the ribozyme gene is under control of a tissue specific promoter. Alternatively, it can be caused by the inducible expression of the ribozyme gene if the ribozyme gene is under control of an inducible promoter. By inducible is meant chemically inducible (by a natural or an artificial molecule) or inducible by environmental factors such as light, drought, stress, pathogen attack and other stimuli. Inducible and tissue specific promoters are well known.

The advantage of controlling gene expression this way is that it makes use of a known ribozyme gene, already shown to function properly. Therefore the need to experiment to find the best molecule for inactivating the useful gene is reduced.

An example of a useful gene is the gene A9-PR-Glucanase (Worrall et al, The Plant Cell, 1992, vol. 4, p 759-771). This gene is used to cause artificial male sterility (AMS) in transgenic plants. AMS plants are used as the female parent in the cross to produce hybrid seeds. When the hybrid plant obtained is grown for its fruits, it is often desirable to restore its fertility to guarantee the optimal fruit set. To facilitate the restoration of the AMS caused by the A9-PR-Glucanase gene, the gene can be engineered by fusing to it the artificial target sequence described above so that the gene function is not affected. It is then possible to use a ribozyme against the artificial target as a restorer of the male sterility. The ribozyme gene is introduced in the male parent where it has no effect (since there is no target sequence in

33

that plant). The hybrid plant resulting from the cross between the AMS female plant and the male restorer plant will be fully fertile since the ribozyme brought in by the male parent will inactivate the AMS gene brought in by the female parent. Other AMS gene can be used similarly. The use of the artificial target is not limited to the area of plant improvement. The artificial target could be used with any artificial gene when an efficient inhibition of the gene is necessary in some circumstances. For example, the artificial target could be appended to a therapeutic gene used in a gene therapy treatment in humans or animals. A tissue specific ribozyme gene against the artificial target could then be used to protect some specific tissue against the activity of the therapeutic gene if such activity was deemed undesirable in said tissue. An inducible ribozyme gene against the artificial target could be used to give the clinician the ability to stop the effect of the therapeutic gene if it was necessary, just by inducing the ribozyme gene. Indeed such systems might be absolutely required for certain gene therapy procedures,perhaps wwhere the expression of toxin genes is involved. A synthetic ribozyme against the artificial target could also be used through extraneaous delivery to control activity of the therapeutic gene in the appropriate tissue or at the appropriate time.

Similarly the artificial target could be used in association with genes used in animal improvement. It could also be used in association with genes used in the food areas for example to improve microorganisms used in food production.

Different aspects of the invention are illustrated in the figures :

34

Figure 1: Location of the primers used to amplify the viral sequences. The amino acid sequence used to design the primers is indicated (see figure2, 3, and 4 below).

Figure 2: Conserved Sequence in NIb coding region, used to design primer 2. Viruses are mentioned with the reference of the publication for their partial (WMV2 and ZYMV) or complete (all others) genomic sequence. Position on the published sequence is indicated. The standard one letter amino-acid code is used.

Figure 3: Conserved Sequence in NIb coding region, used to design primer 3.
Notations are as in Figure 2. CP is Coat Protein.

Figure 4: Conserved Sequence at end of NIa coding region, used to design primer 4. Notations are as in Figure 2.

Figure 5: Sequence of the cDNA of the 3' region of WMV2 strain Mar6. Bases 1-10 were added during cloning. Viral sequence starts at base 11 and ends at base 2971 where the polyA tail starts.

Figure 5 bis: Sequence of the cDNA of the 3' region of WMV2 strain Val3. Bases 1-10 were added during cloning. Viral sequence starts at base 11 and ends at base 2970 where the polyA tail starts.

Figure 6: Sequence of the cDNA of the 3' region of ZYMV strain E15. Bases 1-6 were added during cloning. Viral sequence starts at base 7 and ends at base 2922 where the polyA tail starts.

Figure 7 Sequences used to build the short arm ribozyme gene against WMV2.

A :Short arm ribozyme against the minus strand, cloned into BlueScript SK+ (pBPL4741).
Lower case letters are used for the ribozyme catalytic sequence and the bases from polylinkers. Relevant

restriction sites are in **bold letters**. The sequence
complementary to the ribozyme RNA is shown.

B :Short arm ribozyme against the plus strand,
cloned into BlueScript SK+ (pBPL4742). Notations as
above. The sequence complementary to the ribozyme RNA
is shown.

C : Sequence of the cloned 3' UTR of WMV2 Val3
used in the short arm ribozyme gene. Lower case letter
show polylinker regions from BlueScript or the BamHI
cloning site added by PCR. Cloning sites destroyed
during cloning are shown between brackets. Numbering
refers to the original sequence determined.

Figure 8: Sequences used to build short arm
ribozymes against ZYMV.

A : First combination against the plus strand,
cloned in BlueScript SK+: pBPL4746. Notations as in
figure 7. The sequence complementary to the ribozyme
RNA is shown.

B : First combination against the minus strand,
cloned in BlueScript SK+: pBPL4747. Notations as above.
The sequence complementary to the ribozyme RNA is
shown.

C : Second combination, against both strands,
cloned in BlueScriptSK+: pBPL4750. Notations as above.
The sequence complementary to the ribozyme RNA is
shown.

D : Third combination, against minus strand,
cloned in BlueScriptSK+: pBPL4751. Notations as above.
The sequence complementary to the ribozyme RNA is
shown.

Figure 9: Sequence of the long arm ribozyme gene
against ZYMV: pBPL4760.

36

Base numbers corresponding to the sequence of the +
strand shown in figure 6 are indicated with a * below
the base.

The first 486 bases are targeting the minus strand
(between base 2383 and base 1961), the junction is
shown by ** under the bases, and the last 568 bases are
targeting the + strand (between base 2388 and base
2871).

Notations as above. The sequence complementary to the
ribozyme RNA is shown.

Figure 10 : In vitro assay of ribozyme from
pBPL4734.

Figure 11 : In vitro assay of ribozyme from
pBPL4736.

Figure 12 : In vitro assay of ribozyme from
pBPL4741.

Figure 13 : In vitro assay of ribozyme from
pBPL4742.

Figure 14 : In vitro assay of ribozyme from
pBPL4750 and pBPL4751.

Figure 15 : Plasmids used in the transient assay.

A : Target plasmids pBPL0184, 0185, 0186 and 0187
contain 190 bp, 210 bp, 170 bp and 180 bp respectively
of minus strand sequence of WMV2 cloned in sites NsiI -
NcoI of pGLFuz (Sawyer and Birch 1996).
Target plasmid pBPL0188 contain 540 bp of the plus
strand of WMV2 cloned into sites NsiI - NcoI of pGLFuz.

B : Ribozyme plasmid pBPL0189 contains a 400 bp
insert cloned into the BglII - XhoI sites of a
derivative of pGLFuz, encoding an octameric ribozyme
against WMV2, targeting sites 696(-), 996(-), 1747(-),
2067(-), 2259(+), 2403(+), 2520(+) and 2720(+).
Ribozyme plasmid pBPL0190 contains a 220 bp insert
cloned into the BglII - ShoI sites of a derivative of

37

pGLFuz, encoding a pentameric ribozyme against the
minus strand of ZYMV, targeting sites 1243, 1743
(equivalent to 1747 in WMV2), 1929, 2442 and 2564.

Figure 16 : Transient LUC activity after co-
bonbardment of either an octanmeric ribozyme (pBPL0189)
or a pentameric ribozyme (pBPL0190) with various
plasmids containing different target regions fused 5'
to the luciferase gene. Each co-bombardment consisted
of 2 µg of ribozyme plasmid (pBPL0189 or pBPL0190) and
2 µg of the target plasmid as indicated. Bars represent
standard errors of the mean of three independent
experiments.

Figure 17 : Construction of pGLFuz vector (Sawyer
and Birch 1996). The 1.8 Kb PCR modified uidA gene was
inserted as a NsiI - NcoI DNA fragment upstream of the
luc gene of pZorz to yield pGLfuzTAG. The luc gene of
pGLfuzTAG was removed (by digestion with NotI,
treatment with T4 DNA polymerase, then digestion with
XhoI) and replaced by the 1.9 Kb luc gene from
pSPNF+luc (prepared by digestion with KpnI, treatment
with T4 DNA polymerase and digestion with XhoI),
yielding the uidA :luc fusion construct pGLFuz. B,
GglII; Bl, blund end; K, KpnI; Luc, luc gene from
pZorz; Mod. GUS, PCR modified uidA gene; Nc, NcoI; Ne,
NheI; Ni, NsiI; No, NotI; P, osa artificial promoter
region containing 4ocs-Δ35S-Ac leader regions;
pSPNF+Luc, luc gene from pSPNF+luc; T, ocs 3'
polyadenylation region; TAG, stop codon; X, ShoI.

Figure 18 : Sequence of the region of WMV2 strain
Mar6 (top line = (+) strand, bottom line = (-) strand)
around ribozyme cleavage site 1747 in the (-) strand
(indicated in bold type). Underlined sequence
(positions 1704 → 1873 in (-) strand) represents the
full length artificial target sequence which can be

38

linked to a reporter gene in transient ribozyme assays, or to a gene of interest as a ribozyme - regulatable element. Arrow (↑) indicates nucleotide in (-) strand which is immediately 5' of the ribozyme cleavage site. (G/T) at position 1863 in the (-) strand indicates that the naturally occurring G can be replaced with a T to mutate the stop codon at 1864-1862 (TGA). This facilitates in-frame fusion, and can be introduced using appropriate PCR primers.

Figure 19 : shows the DNA sequence of pBPL3817 (long arm ribozyme against WMV2 NIb and CP cistrons). The sequence is written from 5' to 3'. The first base corresponds to base number 279 in figure 5. The last viral base is G, 34 bases before the end of the sequence.The ribozyme catalytic regions are underlined, in italics

Figure 20 : shows the genome organisation of the type member of the potyirus family, PVY (from Chapter 4, « The Potyviridae », Ed. Shukla, D.D., Ward, C.W., Brunt, A.A, 1994 CAB International). P1 : First Protein ; HC-Pro :Helper component ; P3 : Third Protein ; 6K1, 6K2 : additional 6kb proteins ; CI : Cylindrical Inclusion Protein ; NIa : Small Nuclear Inclusion protein ; NIb : large Nuclear Inclusion Protein ; CP : Coat Protein

Figure 21 : shows the determination of the viral epidemiology in Season 2 of field trials (described in Example 10 below). The technique used was ELISA analysis on virus susceptible control, taken randomly. The notations were made over a period of one month at intervals of one to two weeks.

WMV2 = Watermelon Mosaic Virus 2 ;

CMV = Cucumber Mosaic ;

ZYMV = Zucchini Yellow Mosaic Virus ;

39

PRSV = Papaya Ringspot Virus.

EXAMPLES

Molecular Biology techniques used are essentially those described in Current Protocols in Molecular Biology (Ausubel F.M. et al. Eds, 1994-1997, John Wiley & Sons, Inc., New York.) except where otherwise mentioned.

1. Cloning and Sequencing the 3' region of genomic RNA of target viruses.

Target viruses for this work are Watermelon Mosaic Virus 2 (WMV2) and Zucchini Yellow Mosaic Virus (ZYMV).

To design ribozymes against these viruses, a partial sequence of the genome of field isolates of these viruses from the Mediterranean region was determined as follows:

a - Infected plant material was harvested, and total RNA extracted.

b - The RNA was used as the template for a Reverse Transcription - Polymerase Chain Reaction (RT-PCR) using primers specific for the 3' terminal region of the potyviral genome.

c - The PCR products were cloned and sequenced. These three steps are detailed below:

a - Zucchini plants were mechanically inoculated with strains Mar6 or Val3 of WMV2 or strains E9 or E15 of ZYMV. These strains are field isolates from the Mediterranean region. Leaves were harvested and total RNA was extracted.

b - A sequence comparison of all published potyvirus sequences allowed the identification of regions of sequence conservation among potyviruses.

Figure 1 shows the location of relevant conserved sequences identified on the genome of several potyviruses.

Figure 2 shows the conserved sequence identified in the region coding for protein NIb used to design primer 2.

Figure 3 shows the conserved sequence identified in the region coding for protein NIb used to design primer 3.

Figure 4 shows the conserved sequence identified at the end of the region coding for protein Nia, and used to design primer 4.

Reverse transcription was performed using the primer Poly-T-anchor:

Poly-T-anchor: 5' GGGAGCTCGCTAGCGGATCCT$_{17}$

Primer 1, used in the PCR step, is homologous to the anchor region, with extra bases added to facilitate cloning of the PCR products:

Primer 1:   5' AATTGGGAGCTCGCTAGCGGAT

The other primers were designed to match the consensus sequence of the conserved regions identified as described above, and some bases were added to facilitate cloning. The primers were moderately degenerate to take into account polymorphism. Multiple bases at a given position are indicated in the sequence by a / : T/G means that both T and G could be present in the primer at a given position. N means that any base could be present.

Primer 2: 5' AATTGCTAGCAGTT/GGTT/GGACAATACACTT/AATGG

Primer 3: 5' AATTCTAGAATA/CAT/GATCATCACCATTG/TGC

Primer 4: 5' AATTCTAGATTCTGGAT/AACATTGGATAA/TC/GNAC

c - PCR products of the predicted size were obtained with primers pairs 1-2 (1.7 kb) and 3-4 (1.3

41

kb). The PCR products were purified on gel, cut with the appropriate restriction enzymes, cloned into plasmid BlueScript SK+ (commercially available from Stratagene) and the complete nucleotide sequence of the viral fragment was determined using standard methods known in the art. Most of the sequence was determined from two independent PCR products from each RNA template. For WMV2, a PCR product was also obtained with primer pair 1-4. It was used subsequently to build long arm ribozyme genes (see Example 3 below).

The sequence of the 3' region of WMV2 strain Mar6 is described in Figure 5 and the sequence of the 3' region of WMV2 strain Val3 is described in figure 5 bis. The two strains are more than 99% identical.

The sequence of the 3' region of ZYMV strain E15 is described in Figure 6. The E9 strain was more than 99% identical to E15.

2. Design of ribozymes against WMV2

Both the plus (+, coding, or genomic) RNA strand of the virus and the minus (-, replication intermediate) RNA strand of the virus were targeted by ribozymes.

The following criteria were used when choosing cleavage sites on the target RNA:

Triplets known to be well cleaved in vitro were preferred (Perriman et al., 1992):
GUC, GUA, GUU, UUC, UUU.

Triplets where the codons had limited degeneracy were preferred.

This increases the chance that the target sequence will not evolve easily so that the virus overcomes the resistance. Some examples are shown below

42

for illustration, with the cleaved triplet in **Bold**.
Many more combinations are possible.

For the plus strand:

Aminoacid sequence:	MetSer	MetTyr	MetPhe	Phe
Nucleotide sequence:	AUGUCN	AUGUAU/C	AUGUUU/C	UUU/C

For the minus strand:

Aminoacid sequence (+):	MetLys	MetThr	GluLys	Tyr
Nucleotide sequence (+):	5'AUGAAA/G	AUGACN	GAA/GAAA/G	UAU/C
Nucleotide sequence (-):	3'UACUUU/C	UACUGN	CUU/CUUU/C	AUA/G

Triplets located in a region of sequence
conservation were preferred. This would increase the
chance that the ribozyme could target multiple strains
of the virus or even multiple species of related
viruses. It would also decrease the chance that a virus
resistant to cleavage arises.

Four sites were selected on the plus strand and
four sites were selected on the minus strand. The
number used to describe a given ribozyme refers to the
base <u>underlined</u>, on the + strand in the following
tables, according to the numbering of the sequence
presented in figure 5. Cleavage site is in **Bold** and the
cleaved bond is indicated by a -.

Sites selected on the minus strand:

Number	Target sequence Aminoacid sequence		+ strand sequence
696	5' UUC - A 3'	5' UGAA 3'	Met Lys
996	5' GUC - A 3'	5' UGAC 3'	Met Thr
1747	5' GUA - U 3'	5' AUAC 3'	Tyr
2067	5' UUU - G 3'	5' CAAA 3'	Thr Lys

Sites selected on the plus strand:

43

Number	Target sequence	Aminoacid sequence
2259	5' UUU - A 3'	Phe
2403	5' UUC - U 3'	Phe
2520	5' UUU - U 3'	Phe
2720	5' GUA - A 3'	Gln Och

Of these sites, site 2259 was conserved in ZYMV (at base 2249 on figure 6).

Site 1747 was highly conserved among most potyviruses: a comparison of a 23 base segment surrounding the site shows the following identity:

ZYMV: 23/23

Potato Virus Y: 20/22

Sugar Cane Mosaic Virus: 19/23

Pea Seed-borne Mosaic Virus: 20/23

Plum Pox Virus: 20/23

Maize Dwarf Mosaic Virus: 20/23

The structure and properties of the ribozymes exemplifying the present invention are indicated below in Table 8.

TABLE 8

Construct					STRUCTURE OF RIBOZYME		3'UTR	Catalytic region	IN VIVO RESISTANCE Same strain	IN VIVO RESISTANCE Different strain
pBPL 4738 (pBPL 3817)	WMV2	NIb+CP	(-)	4	long	-	No	Variant (I + II)	Resistance (heterozygous)	Not tested
pBPL 4739 (pBPL 380010)	WMV2	CP+3NT	(+)	4	long	-	No	Variant (I + III)	Resistance (heterozygous)	Not tested
pBPL 4743 (pBPL 3754)	WMV2	NIb+CP+3NT	(+) and (-)	8	short	not contiguous	Yes	Variant (IV)	Resistance (homozygous)	Not tested
pBPL 4748 (pBPL 380?)	ZYMV (E15)	NIa+CP+NIb	(+) and (-)	10	short	not contiguous	Yes	Variant (IV)	E-15 No resistance	13-18 Resistance (homozygous)
pBPL 4750 (pBPL 4?)	ZYMV (E15)	CP	(+) and (-)	5	short	some contiguous	No	variant (IV)	E-15 No resistance	13-18 Resistance (homozygous)
pBPL 4751 (pBPL 4194)	ZYMV (E15)	CP	(-)	2	short	contiguous	No	Variant (IV)	E-15 No resistance	13-18 Resistance (homozygous)
pBPL 4760 (pBPL 4771)	ZYMV (E15)	CP	(+) and (-)	7	short and long	contiguous	No	Natural	E-15 Resistance (heterozygous)	13-18 Resistance (homozygous)

[1] Construct name in brackets indicates designation of ribozyme in plant transformation vector

[2] « 3'UTR » signifies presence, on 3' end of ribozyme construct, of the 3' untranslated region of WMV2 strain Val3 (approximately 300 bases)

[3] Catalytic regions : « natural » signifies sTobRV catalytic domain , « variant » signifies one or more of the catalytic regions comprise mutations in the ribozyme loop, roman numerotation in brackets indicate the type of mutation, as defined above in description

45

3. Construction of a long arm multimeric ribozyme gene against the minus strand of WMV2 (pBPL 4738)

A subclone of the cDNA for the 3' region of WMV2 Mar6, extending from base 269 (EcoRI site) to base 216 8 (XhoI site) (see figure 5) was built by internal deletions of a plasmid containing the PCR product obtained with primers 1 and 4 as described in 1. above. This subclonewas used as a template for the next step. Inverse PCR (IPCR) reactions were used to insert a ribozyme catalytic sequence in replacement of the third base of each target triplet, essentially as described in Unit 3.17.3 of Current Protocols in Molecular Biology. The ribozyme catalytic sequence from the satellite of Tobacco Ringspot Virus (sTobRV, Haseloff and Gerlach, 1988) was used.

The primers used are listed below.

Bases in lower case are encoding the ribozyme catalytic sequence. Bases in upper case are complementary to the target region and used to prime the plasmid template. Bases paired with the first 2 bases of the cleaved triplet are in bold. For the first primer pair the sequence of the final product around site 696 is shown, to demonstrate that ligation of the IPCR product will produce a molecule with a functional catalytic sequence inserted in replacement of the third base of the target site.

Rbz696-3':   5'gtgaggacgaaAAAGCTGCTGTCGGTGCTC3'
Rbz696-5':   3'CTAAATACCTAGGGACTTGTAgactactcagg5'

Result for 696: 5'...GAACAUcugaugaguccgugaggacgaaAAAGC...3'

Rbz996-3':   5'gtgaggacgaaACTAAGTTTTATGGGGGTTGG3'

46

Rbz996-5':   3'CAGGTACCTGTCAACCGTAgactactcagg5'

Rbz1747-3':   5'gtgaggacgaaACATAGCTGAGACAGCATTG
Rbz1747-5':   3'CGTAGGCCTTTTCGAGGTgactactcagg5'

Rbz2067-3':   5'gtgaggacgaaAAGAAAATGAACCTTCCAACAG3'
Rbz2067-5':   3'GGGTGCTGATGTTTTCTATTGgactactcagg5'

The plasmid template was amplified by IPCR using the first pair of primers. The full length PCR product was purified on a gel, and circularised by ligation. The circular product was transformed into bacteria, and the resulting plasmid was extracted from a bacterial colony and submitted to the second IPCR reaction using the second pair of primers. The procedure was repeated with the third and fourth pair, and the final product was analysed by sequencing the modified regions. A plasmid called pBPL4734 was obtained; it had the correct sequence except for ribozyme 696 which had one extra base in the loop, and ribozyme 996, where a small deletion had taken place. Attempts were made to repair errors in 996 by a new round of IPCR. The final product was called pBPL4738 and had the expected sequence except at sites 696 and 996 where the catalytic sequence is mutated in the ribozyme loop (upper case):

Mutant ribozyme 696 :   cugaugaguccCGUGAggacgaa
Mutant ribozyme 996 :   cugaugaguccUGAggacgaa
Wild type sTobRV: cugaugaguccGUGAggacgaa.

To build a functional ribozyme gene in a plant transformation vector, the multimeric long arm ribozyme sequence was cloned in the plant expression cassette from pBIOS512, made of the enhanced 35S promoter and

47

the nos terminator, and the gene was then inserted in the plant transformation vector pSCK.

Description of pBIOS512: this plasmid was built by replacing the 35S promoter region of plasmid pBIOS3 (Perez et al., 1989) by the enhanced 35S promoter region from pJIT163, a plasmid derived from pJIT60 (described in Guerineau and Mullineaux, 1993) by inserting an NcoI site in the polylinker. The enhanced 35S promoter was isolated from pJIT163 as a SstI - BamHI fragment and cloned into sites SstI - BamHI of pBIOS3 in replacement of the 35S promoter.

Description of pSCK: this plasmid was built by inserting a kanamycin resistance gene into the plant transformation vector pSCV1 (Firek et al., 1993). The kanamycin resistance gene was built by inserting the nptII coding sequence from pCaMVNEO (Fromm et al., 1986) as a BamHI fragment in the BamHI site of pBIOS1 (Perez et al., 1989), resulting in pBIOS1K. The gene was then isolated from pBIOS1K as a EcoRI(Klenow) fragment and inserted into site EcoRV of pSCV1, resulting in pSCK.

The ribozyme sequence was cloned in the expression cassette of pBIOS512, as a EcoRV-SmaI - HindIII fragment in the HindIII (Klenow) site of pBIOS512. The gene was then excised as a EcoICRI - EcoRI(Klenow) fragment and inserted into the SmaI site of pSCK.

The resulting plasmid was called pBPL3817. It contains a T-DNA with a marker gene conferring resistance to kanamycin and the long arm ribozyme gene expressing an RNA with approximately 1900 bases complementary to the minus strand of the virus (the target RNA) and targeting four sites on the target RNA.

48

4. Construction of a long arm multimeric ribozyme gene against the plus strand of WMV2 (pBPL 4739)

A procedure similar to that described in 3. was used to build a ribozyme gene against the + strand. The plasmid template used in the IPCR reactions extended from base 2164 (site XhoI) to the end of the cDNA (base 2988, including 18 A of the polyA tail) and was obtained by internal deletions of a partial cDNA clone of the 3' region of WMV2.

Inverse PCR reactions were used as described above in 3. using the following primer pairs:

(same notations as above)

Rbz2259+3':   5'gtgaggacgaaAAACCATTCATAATCACACCC3'
Rbz2259+5': 3'CAATAGCTATGTGGTTTGGTAgactactcagg5'

Rbz2403+3':   5'gtgaggacgaaAAATGGTGCATGATTTGTC3'
Rbz2403+5': 3'CGAAGACGACGCAGACTgactactcagg5'

Rbz2520+3':   5'gtgaggacgaaAAGTCAAAAGCATAGCGTGC3'
Rbz2520+5': 3'CACAAAACCTTCATTGGAGTATgactactcagg5'

Rbz2720+3':   5'gtgaggacgaaACTGCGGTGGACCCATACC3'
Rbz2720+5': 3'CACTGGTCAAATGGATCAGAAgactactcagg5'

The plasmid obtained was called pBPL4736, and had half of ribozyme 2259 missing, and a one base deletion in the 3 other ribozymes' loop. New rounds of IPCR inserted a correct ribozyme for 2259 and corrected ribozyme 2403. The final product still contained some errors which were not considered serious problems. The ribozyme sequence for sites 2259 and 2403 is correct (wild type as described). The ribozyme sequences for sites 2520 and 2720 each have one base deleted:

49

Mutant ribozyme in 2520:   cugaugaguccUGAggacgaa
Mutant ribozyme in 2720:   cugaugagucGUGAGgacgaa
Wild type sTobRV: cugaugaguccGUGAggacgaa

These mutations are known not to abolish ribozyme activity. This was confirmed by in vitro assays (see below). This plasmid was called pBPL4739.

The sequence from pBPL4739 was cloned in the expression cassette of pBIOS512 as a BamHI(Klenow)-EcoRI fragment in the HindIII(Klenow) site of pBIOS512. The gene was then transferred to the plant transformation vector pSCK as described above in 3.

The resulting plasmid was called pBPL38010. It contains a T-DNA with a marker gene conferring resistance to kanamycin and the ribozyme gene expressing an RNA with approximately 830 bases complementary to the plus strand of the virus (the target RNA) and targeting four sites on the target RNA.

5. Construction of a short arm multimeric ribozyme gene against both strands of WMV2(pBPL 4743)

In order to build a short arm multimeric ribozyme targeting the sites described above for the long arm ribozymes, synthetic oligonucleotides encompassing the entire sequence of the desired ribozyme gene were purchased. The oligonucleotides were assembled and the desired full length product was obtained using Ligation based PCR (LbPCR) as described by Betzner et al. (1997). The short arm multimeric ribozyme genes against the minus strand and the plus strand were synthesised independently and assembled using restriction sites built at the end of the synthetic sequences.

The ribozyme catalytic sequence used for these constructs was a variant of the sTobRV ribozyme described above, with a 3 base-pair helix II instead of the natural 4 base-pair. The ribozyme used for the short multimeric ribozyme was :

cugaugaGCGgugaCGCgaa.

(Helix II shown here in upper case letters)

The pairing regions between the ribozymes and the target (helix I and helix III) was made of 2 arms of 10 to 12 bases (see Haseloff and Gerlach, 1988 for a description of helices I, II, and III).

The sequence of the oligonucleotides used is given below; lower case letters are the ribozyme catalytic sequence, lower italic letters are linker regions, upper case letters are target sequences. All sequences are shown 5' to 3'. The sequence synthesised (shown) is actually the strand complementary to the ribozyme RNA.

Oligonucleotides used for the ribozymes against minus strand:

WM2067S :

*gcagctgctggatccatg*TCATTTTCTTttcgcgtcaccgctcatcagGTTATCT TTTG

WM1747S :

TCAGCTATGTttcgcgtcaccgctcatcagTGGAGCTTTTC

WM996S :

*Cagatct*ATAAAACTTAGTttcgcgtcaccgctcatcagATGCCAACTG

WM696S :

CAGCAGCTTTttcgcgtcaccgctcatcagATGTTCAGGGA*cctgcatctgatca agcttg*

51

Guide oligonucleotides used to align the full length sequence for the ligation reaction:

WMGU4S: assemble WM2067S with WM1747S
GAAACATAGCTGACAAAAGATAACCTG

WMGU5S: assemble WM1747S with WM996S
GTTTTATAGATCTGGAAAAGCTCCAC

WMGU6S: assemble WM996S with WM696S
GAAAAAGCTGCTGCAGTTGGCATC

Oligonucleotides used for the ribozymes against plus strand:

WM2259S:
*gcagctgctggatcc*ATGAATGGTTTttcgcgtcaccgctcatcagATGGTTTGG

WM2403S:
TGCACCATTTttcgcgtcaccgctcatcagTCAGACGCAG*at*

WM2520S:
CTTTTGACTTttcgcgtcaccgctcatcagTATGAGGTTAC

WM2720S:
CACCGCAGTttcgcgtcaccgctcatcagAAGACTAGGTAA*cctgcatctgatca*
agcttg

Guide oligonucleotides used to align the full length sequence for the ligation reaction:

WMGU1S: assemble WM2259S with WM2403S
GAAAAATGGTGCACCAAACCATCTG

52

WMGU2S: assemble WM2403S with WM2520S
CGAAAAGTCAAAAGATCTGCGTCTGAC

WMGU3S: assemble WM2520S with WM2720S
AAACTGCGGTGGTAACCTCATACTG

Primers used to amplify the assembled ligated
oligonucleotides by PCR:

WMPC1S: AATTGCAGCTGCTGGATCCATG
WMPC2S: AATTCAAGCTTGATCAGATGCAG

The PCR products were digested with BamHI and
HindIII and cloned into BlueScript SK+. The ribozyme
against the minus strand gave plasmid pBPL4741, and the
ribozyme against the plus strand gave pBPL4742. The
sequences of the relevant regions are on figure 7.
Restriction sites used for further cloning are
indicated.

The 2 ribozymes were combined by inserting the
BstXI - BspMI fragment of pBPL4742 in sites BstXI -
BamHI of pBPL4741, yielding pBPL4743.

The 3' non coding region of WMV2 strain Val3 (3'
UTR) was then cloned in 3' of the multimeric ribozyme
sequence. This was done as an attempt to increase the
efficacy of the ribozyme by co-compartmentalisation
with the target viral RNA. If 3' end sequences of RNA
are involved in cellular localisation of RNA, as is the
case with many RNAs, it was hoped that by providing the
ribozyme RNA with a 3' sequence similar to that of its
target, the 2 molecules would be present in the same
cellular compartment.

The 3' UTR region of WMV2 Val3 was cloned from
the original cDNA (see sequence in figure 5 bis) as a
HinfI (at base 2678) - BamHI (from the anchor region in

primer 1, see 1. above) fragment into EcoRI - BamHI of BlueScript SK+. The fragment contains bases 2678 to 2970 of strain Val3 cDNA, which contains the last 14 codons of the cDNA followed by the complete untranslated region. This sequence (shown in figure 7) was then inserted as an EcoRV - NotI fragment into BamHI(Klenow) - NotI sites of pBPL4743, giving pBPL4744.

The now complete sequence was then transferred as a BamHI - Hind III fragment in sites BamHI - HindIII of pBIOS512 to create a functional gene, as described above in 3. This gene was then cloned in the SmaI site of pSCK as described above in 3., yielding pBPL3754.

pBPL3754 contains a T-DNA with a marker gene conferring resistance to kanamycin and the ribozyme gene expressing an RNA approximately 700 bases long. This RNA contains 8 regions of homology to the plus or minus strand of the virus. Each region is between 20 and 22 bases long, and is split in the middle by the ribozyme catalytic sequence. In addition the ribozyme RNA has a 3' region containing approximately 300 bases identical to the viral 3' end.

6. Construction of several short arm multimeric ribozyme genes against ZYMV

Site selection will be described using the same notations as in 2. above. The numbering refers to the sequence shown on figure 6.

Sites selected on the minus strand:

Number	Target sequence	+ strand sequence	Aminoacid sequence
1245	5' GUU - G 3'	5' CAAC 3'	Asn

54

1743	5' GUA – U 3'	5' AUAC 3'	Tyr
1929	5' GUC – A 3'	5' UGAC 3'	Asp
2009	5' UUC – A 3'	5' UGAA 3'	ValAsn
2027	5' UUC – C 3'	5' GGAA 3'	GlyLys
2354	5' UUU – U 3'	5' AAAA 3'	GluAsn
2442	5' GUA – U 3'	5' AUAC 3'	Tyr
2564	5' UUC – A 3'	5' UGAA 3'	MetLys

Sites selected on the plus strand:

Number	Target sequence	Aminoacid sequence
110	5' GUC – U 3'	Val
2161	5' GUC – U 3'	AlaSer
2255	5' GUU – U 3'	Val
2455	5' GUA – U 3'	ArgTyr
2466	5' UUC – G 3'	LeuArg
2590	5' UUC – U 3'	ValSer
2603	5' UUU – G 3'	Phe

Several combinations of short arm ribozymes were built using synthetic oligonucleotides. The procedure was similar to the procedure described in 5. above.

6.1. Construction of pBPL 4748

The first combination contained 10 ribozyme activities, 5 against each strand:
1245 – 1743 – 1929 – 2442 – 2564 against the minus strand (plasmid pBPL4747)
110 – 2161 – 2255 – 2455 – 2590 against the plus strand (plasmid pBPL4746)

55

The ribozyme catalytic sequence used for these constructs was the same variant of the sTobRV ribozyme as described in 5. above, with a 3 base-pair GCG helix II instead of the natural 4 base-pair.

The pairing regions between the ribozymes and the target (helix I and helix III) was made of 2 arms of between 8 and 11 bases.

Sequence of the relevant region of pBPL4746 and pBPL4747 is shown on figure 8.

The sequences were assembled together with the 3'UTR of WMV2 Val3 (shown on figure 7), by a tripartite ligation: the EcoRV(de phosphorylated) – NotI fragment containing the 3'UTR and the BamHI(Klenow) – HindIII(Klenow) fragment from pBPL4747 were ligated together in pBPL4746 cut with BamHI(Klenow) and NotI (both sites de phosphorylated). The resulting plasmid was called pBPL4748 and contained (from HindIII to BamHI) the sequence for the ribozymes against the plus strand, followed by the ribozymes against the minus strand followed by the 3'UTR. The complete sequence was then cloned as a BamHI – HindIII fragment in the expression cassette of pBIOS512 as described in 3. above, and the functional gene was then cloned into the plant transformation vector pSCK as described in 3. above yielding pBPL3906.

Plasmid pBPL3906 contains a T-DNA with a marker gene conferring resistance to kanamycin and the ribozyme gene expressing an RNA approximately 750 bases long. This RNA contains 10 regions of homology to the plus or minus strand of ZYMV. Each region is between 16 and 20 bases long, and is split in the middle by the ribozyme catalytic sequence. In addition the ribozyme RNA has a 3' region containing approximately 300 bases identical to the non-translated 3' end of WMV2 Val3.

56

6.2 Construction of pBPL4750

A second combination was made using the following
ribozymes:
2354 (-) - 2455 (+) - 2466 (+) - 2590 (+) - 2603 (+).
The synthetic sequence cloned into BlueScript SK+
yielded pBPL4750.

This was an attempt to capitalise from the
clustering of sites 2455-2466 and 2590-2603.

The ribozyme catalytic sequence used for these
construct was the natural sTobRV hammerhead ribozyme
except for 2354 and 2603 where the same variant of the
sTobRV ribozyme as described in 5. above, with a 3
base-pair GCG helix II instead of the natural 4 base-
pair was used.

The pairing regions between the ribozymes and the
target (helix I and helix III) was made of 2 arms of
between 8 and 12 bases. Adjacent ribozymes 2455-2466
and 2590-2603 had one arm in common.

The relevant sequence of pBPL4750 is shown on
figure 8. The ribozyme sequence was cloned directly as
a BamHI - HindIII fragment in the expression cassette
of pBIOS512 as described in 3. above, and the
functional gene was then cloned into the plant
transformation vector pSCK as described in 3. above
yielding pBPL4186.

Plasmid pBPL4186 contains a T-DNA with a marker
gene conferring resistance to kanamycin and the
ribozyme gene expressing an RNA approximately 250 bases
long. This RNA contains 3 regions of homology to the
plus and minus strand of ZYMV. One region of 21 bases
is split in the middle by one ribozyme catalytic

57

sequence. The other 2 regions of 31 and 32 bases are each split in three parts by two ribozyme sequences.

6.3. Construction of pBPL 4751

A third combination was made using a double ribozyme against sites 2009 - 2027 (both minus strand). The synthetic sequence cloned into BlueScript SK+ yielded pBPL4751.

The ribozyme catalytic sequence used for this construct was the same variant of the sTobRV ribozyme as described in 5. above, with a 3 base-pair GCG helix II instead of the natural 4 base-pair.

The pairing regions between the ribozymes and the target (helix I and helix III) was made of 2 arms of 10 and 17 bases for each ribozyme. The 17 base arm was shared between the 2 ribozymes. In effect the total pairing region was 37 bases (10 + 17 + 10), interrupted in 2 places by the 2 ribozyme sequences.

The relevant sequence of pBPL4751 is shown on figure 8. The ribozyme sequence was cloned directly as a BamHI - HindIII fragment in the expression cassette of pBIOS512 as described in 3. above, and the functional gene was then cloned into the plant transformation vector pSCK as described in 3. above yielding pBPL4194.

Plasmid pBPL4194 contains a T-DNA with a marker gene conferring resistance to kanamycin and the ribozyme gene expressing an RNA approximately 100 bases long. This RNA contains 1 region of homology to the minus strand of ZYMV. The region is 37 bases long, and is split in three parts by the 2 ribozyme catalytic sequences.

58

7. Construction of a long arm multimeric ribozyme gene against ZYMV

A single gene, expressing one polyribozyme RNA targeting several sites on both strands was made.

The sites chosen were (see 6. above for numbering) :
2009 (-) , 2027 (-) , 2354 (-) , 2455 (+) , 2466 (+) , 2590 (+) and 2603 (+).

To create the appropriate template, two fragments from the cDNA of the 3' region of ZYMV were cloned in opposite orientation as follows (see figure 6 for sequence).

The cDNA clone obtained from ZYMV E15 with PCR primers 1 and 2 (as described in 1. above) was cut with NsiI (at base 2383) and SacII (in the polylinker, next to 5'end of the cDNA), then treated with Klenow and de phosphorylated. The large fragment was then purified and used as a vector in the subsequent ligation. The same plasmid was cut with NsiI (at base 2383) and AlwNI (at base 1955), then treated with Klenow, and the 470 base fragment was purified and used as the insert in the subsequent ligation. The clones obtained were screened for the appropriate orientation, and the selected plasmid was called pBPL0157. This plasmid contains the sequence between bases 2383 and 1961 linked to the sequence between bases 2388 and 2922. The ligation places side by side the region 3' of 1961 and the region 3' of 2383. When a promoter is placed in the appropriate orientation, the transcription product of this sequence will be complementary to the minus strand of ZYMV between base 1961 and 2383, and complementary to the plus strand of ZYMV between base 2388 and 2922.

59

Ribozyme catalytic sequences were inserted at the positions mentioned above using a sequence of IPCR reactions on template plasmid pBPL0157, following the procedure described for WMV2 in 2. above. The ribozyme catalytic sequence used for this construct was the natural sTobRV hammerhead ribozyme as described previously.

The insertion of the catalytic sequence targeting site 2354 (-) proved difficult, as if this ribozyme sequence was toxic to E. coli or affected plasmid replication.

The plasmid containing the 7 ribozyme sequences was called pBPL4760. Sequence of the ribozyme gene in pBPL4760 is shown in figure 9.

The complete sequence was recovered using HindIII (which cleaves at base 2866) and SacI(Klenow) (from the polylinker) and the fragment was cloned in the expression cassette of pBIOS512 in sites HindIII - SalI(Klenow). This had the effect to remove from the region targeting the plus strand the last 51 nucleotides. The functional gene was then transferred to the plant transformation vector pSCK as described previously, yielding plasmid pBPL4771.

Plasmid pBPL4771 contains a T-DNA with a marker gene conferring resistance to kanamycin and the long arm ribozyme gene expressing an RNA with approximately 900 bases complementary to the plus and minus strand of ZYMV RNA and targeting seven sites on the target RNAs.

8. In vitro assay of the ribozymes

Ribozyme activity was assayed in vitro by incubating an appropriate target RNA, prepared by in vitro transcription and uniformly labelled with

radioactive UTP, with the ribozyme RNA prepared by in vitro transcription. The protocol is described in detail below and the result for each ribozyme are also shown.

Target RNA preparation.

The plasmids used for target RNA synthesis were all derived from the plasmids containing the original PCR amplified viral sequences described in 1. above. The T3 and T7 promoters present in the BlueScript vector were used to transcribe the plus or the minus strand of the cloned fragment as appropriate. In vitro transcription using T3 RNA Polymerase or T7 RNA Polymerase from Promega or New England Biolabs was done following the suppliers instructions (for example in Promega Protocols and Application Guide, available from the Promega company, Madison, Wisconsin, USA), in the presence of radioactive α^{32}P-UTP. The template plasmid was linearised with an appropriate restriction enzyme prior to transcription. The target RNA was quantitated by measuring the incorporation of radiolabel using the DE81 filter binding assay as described in the Promega Protocols. Target RNA was stored as a precipitate under ethanol at -20°C, and resuspended in water at between 0.01 and 0.1 pmol/µl when needed. Specific activity was usually between 20,000 and 200,000 cpm/pmol.

Ribozyme RNA preparation

The plasmids described above in sections 3, 4, 5, 6 and 7, made of the various polyribozyme sequences cloned into BlueScript, were used as templates for in vitro transcription reactions as described above, but using only non radioactive nucleotides. A pilot reaction with a small amount of radioactive tracer was done separately to quantitate the synthesised RNA,

61

using the DE81 filter binding assay. Ribozyme RNA was
stored as a precipitate under ethanol at -20°C and
resuspended in water at between 0.1 and 1 pmol/µl when
needed.

Ribozyme reaction

        The target RNA and ribozyme RNA were mixed in T7
RNA Polymerase buffer (from New Englands Biolabs) and
incubated at various temperature (22°C, 37°C or 50°C).
The standard reaction was in 4 µl and contained 0.005
to 0.1 pmol of target, and between 1 and 10 fold molar
excess of ribozyme. The reaction was stopped by adding
an equal volume of stop-loading buffer (Formamide
containing 10mM EDTA and 1mg/ml each Xylene Cyanol and
Bromophenol Blue) to the reaction and freezing the
sample at -20°C until all samples were ready. The
reactions were then boiled for 30 seconds, placed on
ice, and loaded on a denaturing polyacrylamide gel in
TBE buffer with 7M urea. Polyacrylamide concentration
was between 4 and 6 % depending on the expected length
of the products. The gel was transferred to a Whatmann
3MM paper, dried and exposed to a PhosphorImager
screen. An autoradiogram of the gel was obtained by
scanning the screen using a PhosphorImager.

Long arm ribozyme against the minus strand of WMV2

        The in vitro assays of long arm ribozymes proved
difficult since the target and the ribozymes do not
readily dissociate prior to loading and/or reassociate
in the gel. The data are therefore of poor quality and
difficult to interprete.

        The ribozyme used was obtained from pBPL4734
(described in 3. above). The target contained 1898
bases of viral sequence, complementary to the sequence

from base 2168 (site XhoI) to base 270 (site EcoRI) of
WMV2 Mar6, with some sequence from the polylinker of
BlueScript KS+ (between the T3 promoter and the XhoI
site).

Figure 10 shows the result of one assay. There
was 4 fold excess ribozyme and the reaction was done at
50°C for 30, 60, and 90 mn. Products clearly
identifiable include the fragments generated by
cleaving all four sites (shown by arrows on the gel).
However it seems that sites 996 and 1747 were
preferentially cleaved.

Long arm ribozyme against the plus strand of WMV2

The ribozyme used was obtained from pBPL4736
(described in 4. above). The target contained 807 bases
of viral sequence, between base 2163 (site XhoI) and
base 2970 (start of polyA tail), with the polyA tail
and polylinker sequence from BlueScript KS+ (between
the T7 promoter and the EcoRI site).

Figure 11 shows the result of one assay.
Incubation of target with 4 fold excess ribozyme at
three different temperatures showed that site 2259 was
not cleaved as expected because of the mutation in that
ribozyme (see 4. above). Site 2720 was the fastest
cleaver and gave products at 22°C. Site 2403 was next
and site 2520 appears the slowest.

When the ribozyme from plasmid pBPL4739 was
tested, evidence of cleavage at site 2259 was seen
(data not shown).

Short arm ribozyme against the minus strand of WMV2

The ribozyme used was obtained from pBPL4741
(described in 5. above). The target contained 1880
bases of viral sequence complementary to the sequence

63

bewteen base 2278 (ClaI site) and base 398 (SpeI site),
with polylinker sequence in 5' from BlueScript KS+
(betwen the T3 promoter and the ClaI site).

Figure 12 shows the result of an assay. All four
sites are cleaved at high temperature. At 22°C, all
sites except 696 are cleaved but cleavage is very slow
as demonstrated by the very low abundance of double
cleaved products.

Short arm ribozyme against the plus strand of WMV2

The ribozyme used was obtained from pBPL4742
(described in 5. above). The target was the same as
described above for the long arm ribozyme.

Figure 13 shows the result of an assay. Note that
the molar excess ribozyme over the substrate was 4.5
(+), 9 (++) and 13.5 (+++). All four sites are cleaved
at 50°C; with four fold excess ribozyme, cleavage is
essentially complete after 1.0 to 1.5 hour. At 22°C,
only sites 2259 and 2403 are cleaved, with ribozyme
2259 being the fastest cleaver.

Short arm ribozyme against ZYMV: first combination
(pBPL4746 and pBPL4747)

The 2 polyribozymes showed some cleavage activity
at 50°C, but very little at 37°C and none detectable at
22°C.

At 50°C, for the minus strand, all sites were
cleaved; site 2442, 1743, and 1245 were cleaved better
than 2564 and 1929.

For the plus strand, all sites were cleaved;
sites 110, 2590 and 2255 were cleaved faster than 2455
and 2161.

64

Short arm ribozyme against ZYMV: second combination (pBPL4750)

The polyribozyme showed cleavage activity on the plus strand at all 4 predicted sites only at 50°C. No activity was detected at 22°C. The same polyribozyme however showed good cleavage activity on the only site present on the minus strand (site 2354) both at 50°C and 22°C.

Figure 14 shows the result of an assay of this ribozyme on the minus strand target. The target contained 1058 bases of viral sequence, complementary to the sequence between base 2633 (SpeI site) and base 1575 (XbaI site) as shown on figure 6, with 39 bases added in 5' (polylinker sequence between the T3 promoter and the XbaI site of BlueScript SK+) and 11 bases added in 3' (polylinker sequence between the XbaI and the BamHI site of SK+; BamHI was used to linearise the plasmid prior to in vitro transcription). For this experiment the labelled ribozyme from the pilot reaction used to quantitate was mixed with the unlabelled ribozyme. The position of the ribozyme band on the gel is shown by an arrow. The ribozyme was used at 5 fold molar excess over the substrate.

Short arm ribozyme against ZYMV: third combination (pBPL4751)

The ribozyme from pBPL4751 (targeting sites 2009 and 2027 on the minus strand) was assayed on the same target as the ribozyme from pBPL4750.

Figure 14 shows the result of this assay. Both sites were cleaved efficiently at both 22°C and 50°C. The presence of a two bands of equal intensity for the products 5' - 2 and 5' - 3 (see figure 14) suggest that

65

site 2009 is cleaved more readily or faster than 2027:
about 50% of cleaved products are cleaved only at 2009.

9. Determination of resistance in transgenic plants

9.1. Genetic transformation of melon

The method used is described in the following
patent applications : French application n°8910848 of
11/08/89, European application n°904022829 of 09/08/90
and United States patent n° 5,422,259 of 06/06/95.

Genetically transformed plants regenerated,
checked by PCR amplification for the presence of nptII
gene, were transferred into pots containing a mixture
of compost and coarse sand (2:1) and placed in culture
in a greenhouse. During the first seven days, the
plants are protected by a plastic cover in order to
maintain a humid atmosphere. The plants are watered
daily. The plants are induced to flower, followed by
self-fertilization. The seeds (T1 generation) are
collected.

9.2. Artificial inoculation of WMV2 and ZYMV on melon to detect resistant plants

Virus strains :
The strain of WMV2 used is called « MAR ».
Two strains of ZYMV pathotype 0 are used:
      - one called « 13-18 » (pathotype F) is a wilting
strain because of the symptoms induced on susceptible
cultivars which bear the Fn gene (Flaccida necrosis)
(Lecocq and Pitrat, 1984).
      - one called « E15 » (pathotype NF) induces vein
clearing symptoms for all the susceptible cultivars

66

(Lecocq and Pitrat, 1984) independent of the presence or not of the Fn gene.

Virus storage :
Leaves showing symptoms (vein clearing) are thinly cut with a razor blade and placed in boxes with calcium chloride to dry them. The boxes are stored at 6°C during a maximum of one year (Bos, 1969).

Virus multiplication :
Melon plantlets of a susceptible variety at the two leaf stage are inoculated by rubbing the two cotyledons with the virus solution (Marrov, 1967).

Virus solution :
One gram of leaves showing symptoms or one gram of the stored virus is ground in 4 ml of extraction buffer (1.075 g of disodium hydrogen phosphate, added with 0.2 g of sodium diethyldithiocarbamate for 100 ml of distilled water) added with activated charcoal (0.1 g) and carborundum (0.0875 g) (Marrov, 1967). This virus solution has to be kept at cool temperature.

Virus inoculation :
Forty-six to forty-eight seeds (T_1 progeny) per genetic transformation event are sown in 10 cm square pots filled with a potting soil. The temperature is maintained from 18 to 25 °C with a photoperiod of 12 hours. Mechanical inoculation on melon plantlets is carried out at the first leaf stage (approximately 15 days after sowing) (Marrov, 1967). The plantlets are inoculated in the same way as for virus multiplication.

Observation of the symptoms :

The symptoms are scored between 2 and 4 weeks after inoculation.

For WMV2:

- susceptible plant: vein banding

- tolerant plant: light vein banding or symptoms appearing late

- resistant plant: no symptoms

For ZYMV:

- susceptible plant (absence of the Fn gene): vein clearing with the E15 or the 13-18 strains.

- susceptible plant (presence of the Fn gene): vein clearing with E15 strain and wilting reaction with the 13-18 strain.

- resistant plant: no symptoms either with E15 or 13-18 strain.

In some cases, the control plants do not show symptoms of virus attack due to bad inoculation of the cotyledons. The behaviour of these control plants enables the determination of the value of the pathology test. « Escape » plants are considered to have escaped from a « correct » virus inoculation.

9.3. Resistance of transformed plants to WMV2 and ZYMV

The transformed plants are demonstrated to be resistant to WMV2 and to ZYMV.

A. Resistance to WMV2 of melon plants genetically transformed with pBPL3754

68

With construct pBPL3754, eight genetic transformation events were analysed for WMV2 resistance behaviour (see table 1).

For six genetic transformation events, a resistance to WMV2 was observed. The segregation ratio (susceptible:resistant) was 3:1, as expected for a single locus being introduced.

B. Resistance to WMV2 of melon plants genetically transformed with pBPL38010

With construct pBPL38010, fifteen genetic transformation events were analysed for WMV2 resistance behaviour (see table 2).

For twelve genetic transformation events, a resistance to WMV2 was observed. The segregation ratio (susceptible:resistant) was 3:1 for nine genetic transformation events. For one genetic transformation event, the segregation ratio (susceptible:resistant) was 1:3.

C. Resistance to WMV2 of melon plants genetically transformed with pBPL3817

With construct pBPL3817, twenty-one genetic transformation events were analysed for WMV2 resistance behaviour (see table 3).

For fourteen genetic transformation events, the segregation ratio (susceptible:resistant) was 1:3 and for three genetic transformation events, the segregation ratio (susceptible:resistant) was 3:1.

69

D. Resistance to ZYMV of melon plants genetically transformed with pBPL3906

With construct pBPL3906, thirteen genetic transformation events were analysed for ZYMV resistance behaviour (see table 4).

With the strain 13.18, the progeny of ten genetic transformed plants was tested. Four genetic transformation events showed a resistance segregation with a ratio (susceptible:resistant) of 3:1.

With the strain E15, all the progeny tested from nine genetic transformation events were susceptible.

E. Resistance to ZYMV of melon plants genetically transformed with pBPL4186

With construct pBPL4186, five genetic transformation events were analysed for ZYMV resistance behaviour (see table 5).

One genetic transformation event from the five tested showed a segregation resistance (ratio susceptible:resistant of 3:1) when tested with the strain 13.18.

With the strain E15, the progeny from the three independent genetic transformation events tested were susceptible.

F. Resistance to ZYMV of melon plants genetically transformed with pBPL4194

70

With construct pBPL4194, eight genetic transformation events were analysed for ZYMV resistance behaviour (see table 6).

With the strain 13.18, the progeny of four genetic transformation events from the seven tested showed a resistance segregation with a ratio (susceptible:resistant) of 3:1.

With the strain E15, the progeny of the six genetic transformation events tested were susceptible.

G. Resistance to ZYMV of melon plants genetically transformed with pBPL4771

With construct pBPL4771, six genetic transformation events were analysed for ZYMV resistance behaviour (see table 7).

With the strain 13.18, two genetic transformation events were studied. The progeny of one genetic transformation event showed a segregation ratio (susceptible:resistant) of 3:1 and the other one a ratio of 1:3.

With the strain E15, the progeny of two genetic transformation events from the six tested showed a resistance segregation with a ratio (susceptible:resistant) of 1:3.

Example 10 - WMV2 Resistance in the Field

Field trials of genetically transformed melon for virus resistance have been carried out.

10.1 - Season n° 1 :

71

This trial concerned the genetic transformation event W38 (insertion of polyribozyme of pBPL3817 in a melon line called RZ33). Molecular analysis of its genomic DNA showed a single locus integration of the T-DNA.

Twelve varieties were grown :

- variety 1 : generation T1 of the event W38
- variety 2 : generation T2 of the event W38
- variety 3 : progeny obtained from the cross between generation T2 of the event W38 and the untransformed line RZ33
- variety 4 to 11 : F1 hybrids realized with W38 (T3 generation) and other parental lines
- variety 12 : untransformed line, RZ33

There were 3 tests for each variety and for each test, 9 plants were cultured. Plants were grown in accordance with normal agronomical practice and scored three times per season for virus symptoms.

The observations of symptoms were noted 0 to 9 as followed :

- 0 : no virus symptom,
- 5 : severe virus symptoms,
- 9 : the plant is nearly dead, covered by virus symptoms.

Some leaf samples were taken for ELISA analysis. The observation data and ELISA values are set out in Table 9 below.

Non transgenic parent plants and control plants were heavily attacked by WMV2, showing that the infection pressure was high that year in this area. Typical foliar mosaic symptoms were observed on the non-transformed material.

By contrast, at the end of the season, all transgenic plants (varieties 1 to 11) had remained

72

immune to WMV2. This can be seen by comparison of the optical density for the plant under test with those of the safe control x2 and inoculated control.

Legend of Table 9 :
Safe control : non-transformed, non inoculated plant.
Inoculated control : non-transformed, inoculated plant.
% of viral infection : natural infection pressure for the indicated virus.
??? : indicates samping error.
"-" : indicates optical density below that of safe control.
Blank space : indicates no sample taken.

10.2 - Season n° 2 :
    The results obtained from the field trial in Season n° 2 are set out below.
    This trial concerned the genetic transformation event W38 (insertion of polyribozyme of pBPL3817 in a melon line called RZ33).
    Six varieties were grown :
    - variety 1 : untransformed line, RZ33
    - variety 2 : generation T2 of the event W38
    - variety 3 : generation T3 of the event W38
    - variety 4 : generation T4 of the event W38
    - variety 5 : F1 hybrid realized with W38 (T4 generation)
    - variety 6 : control F1 hybrid realized with RZ33

    There were 3 tests for each variety and for each test, 8 plants were cultured. Plants were grown according to normal agronomical practice and scored five times per season for virus symptoms.
    The observations of symptoms were noted 0 to 9 as followed :

73

- 0 : no virus symptom
- 5 : severe virus symptoms
- 9 : the plant is nearly dead, covered by virus symptoms

Some leaf samples were taken for ELISA analysis. The observation data and ELISA values are presented in Table 10 below.

Non transgenic parent plants and control plants were heavily attacked by WMV2 and CMV, showing that the infection pressure was high that year in this area. Figure 21 shows the viral epidemiology in the field trial of Season n° 2, for WMV2, CMV, ZYMV, and PRSV. Infection pressure for WMV2 was high during Season n° 2 (see Table 10 "% of viral infection"). Typical foliar mosaic symptoms were observed on the non-transformed material. The virus symptom development expanded to the fruit as raised green "warts" and would therefore be considered non commercial. The identity of the virus was checked by ELISA.

By contrast, at the end of the season, all transgenic plants (varieties 2 to 5) had remained immune to WMV2. Transfenic fruits were safe and of high quality.

Legend of Table 10
See legend of Table 9.

TABLE 1. RESULTS OF THE RESISTANCE TEST AGAINST WMV2 on T_1 PROGENY FROM PLANTS GENETICALLY TRANSFORMED PLANTS WITH pBPL3754

Name of the T_0 plant	Number of T_1 plants tested	Number of non-scorable plants	Number of susceptible plants	Number of tolerant plants	Number of resistant plants	Most probable genetic determinism of resistance
2.1	46	0	44	2	0	escape?
4.4	45	0	33	6	6	insertion at 1 locus of a resistance gene efficient in the homozygous condition
4.10	43	0	21	8	14	Insertion at 1 locus of a resistance gene efficient in the homozygous condition
4.12	24	0	16	4	4	insertion at 1 locus of a resistance gene efficient in the homozygous condition
8.1	47	0	36	0	11	Insertion at 1 locus of a resistance gene efficient in the homozygous condition
93.16	45	5	37	1	2	escape?
93.42	45	1	34	4	6	insertion at 1 locus of a resistance gene efficient in the homozygous condition
93.47	26	0	19	2	5	insertion at 1 locus of a resistance gene efficient in the homozygous condition

TABLE 2. RESULTS OF THE RESISTANCE TEST AGAINST WMV2 on T₁ PROGENY FROM PLANTS GENETICALLY TRANSFORMED WITH pBPL38010

Name of the T$_0$ plant	Number of T$_1$ plants tested	Number of non-scorable plants	Number of susceptible plants	Number of tolerant plants	Number of resistant plants	Most probable genetic determinism of resistance
43.2	45	0	43	2	0	/
57.1	45	1	43	1	0	/
71.5	33	0	14	0	19	Insertion at one locus ?
71.7	23	1	17	0	5	insertion at 1 locus of a resistance gene efficient in the homozygous condition
73.17	46	0	46	0	0	/
73.20	20	0	16	0	4	insertion at 1 locus of a resistance gene efficient in the homozygous condition
80.9	37	0	28	3	6	insertion at 1 locus of a resistance gene efficient in the homozygous condition
86.9	46	11	16	5	14	insertion at one locus ?
86.21	46	0	34	0	12	Insertion at 1 locus of a resistance gene efficient in the homozygous condition
86.24	12	0	9	0	3	insertion at 1 locus of a resistance gene efficient in the homozygous condition
86.39	13	0	5	2	6	insertion at one locus ? of a resistant gene (germination problem)
86.43	45	0	19	13	13	insertion at 1 locus of a resistance gene efficient in the homozygous condition
92.25	44	1	17	17	9	insertion at 1 locus of a resistance gene efficient in the homozygous condition
92.31	46	1	18	10	17	insertion at 1 locus of a resistance gene efficient in the homozygous condition?
92.86	17	1	5	2	9	Insertion at 1 locus of a resistance gene efficient in the heterozygous condition

TABLE 3. RESULTS OF THE RESISTANCE TEST AGAINST WMV2 on T_1 PROGENY FROM PLANTS GENETICALLY TRANSFORMED WITH pBPL3817

Name of the T_0 plant	Number of T_1 plants tested	Number of non-scorable plants	Number of susceptible plants	Number of tolerant plants	Number of resistant plants	Most probable genetic determinism of resistance
30.1	7	6	1	0	0	look of an haploid plant
38.2	46	0	0	0	46	insertion at multiloci
44.2	18	8	4	1	5	insertion at 1 locus of a resistance gene efficient in the heterozygous condition (germination problem)
46.1	46	7	2	2	35	insertion at 2 loci of a resistance gene efficient in the heterozygous condition
46.2	46	0	45	0	1 escape?	/
46.3	44	2	9	4	29	insertion at 1 locus of a resistance gene efficient in the heterozygous condition
48.1	46	5	7	1	33	insertion at 1 locus of a resistance gene efficient in the heterozygous condition
72.3	46	0	10	11	25	insertion at 1 locus of a resistance gene efficient in the heterozygous condition
72.6	27	0	3	0	24	insertion at 1 or 2 loci of a resistance gene efficient in the heterozygous condition
72.7	36	1	3	5	27	insertion at 1 or 2 loci of a resistance gene efficient in the heterozygous condition
74.11	5	0	0	0	5	insertion at 1 or 2 loci of a resistance gene efficient in the heterozygous condition germination problem

(table 3 - continued)

RESULTS OF THE RESISTANCE TEST AGAINST WMV2 on T_1 PROGENY FROM PLANTS GENETICALLY TRANSFORMED WITH pBPL3817

Name of the T_0 plant	Number of T_1 plants tested	Number of non-scorable plants	Number of susceptible plants	Number of tolerant plants	Number of resistant plants	Most probable genetic determinism of resistance
77.1	41	0	10	7	24	insertion at 1 locus of a resistance gene efficient in the heterozygous condition
77.3	44	1	6	3	34	insertion at 1 locus of a resistance gene efficient in the heterozygous condition
77.4	31	0	2	0	29	insertion at 2 loci of a resistance gene efficient in the heterozygous condition
77.5	44	0	7	0	37	insertion at 1 locus of a resistance gene efficient in the heterozygous condition
77.9	39	0	18	6	15	insertion at 1 locus of a resistance gene efficient in the homozygous condition
77.11	37	0	24	7	6	insertion at 1 locus of a resistance gene efficient in the homozygous condition
85.2	41	0	17	4	20	? ?
87.9	46	0	35	1	10	insertion at 1 locus of a resistance gene efficient in the homozygous condition
87.13	40	0	5	18	17	insertion at 1 or 2 loci of a resistance gene efficient in the heterozygous condition?
87.15	11	0	4	4	3	insertion at 1 locus of a resistant gene

TABLE 4. RESULTS OF THE RESISTANCE TEST AGAINST ZYMV (strains 13.18 FN and E15) on T_1 PROGENY FROM PLANTS GENETICALLY TRANSFORMED WITH pBPL3906

Name of the T_0 plant	Number of T_1 plants tested	Number of non-scorable plants	"FN reaction" or susceptible plants	Number of resistant plants	Most probable genetic determinism of resistance
Strain 13.18, FN					
Z6.3	37	0	37	0	/
Z6.4	43	0	43	0	/
Z6.5	33	0	33	0	/
Z6.6	18	0	18	0	/
Z6.7	15	0	15	0	/
Z11.41	38	0	34	4	escape?
Z11.42	23	0	16	7	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Z19.12	30	0	20	10	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Z19.20	10	0	8	2	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Z19.21	19	0	11	8	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Strain E15					
Z6.3	36	0	36	0	/
Z6.4	45	0	45	0	/
Z6.5	17	0	17	0	/
Z6.6	8	0	8	0	/
Z6.7	10	0	10	0	/
Z19.4	8	0	5	0	/
Z19.20	11	0	11	0	/
Z19.23	22	0	12	1	/
Z20.1	23	12	9	2	/

TABLE 5. RESULTS OF THE RESISTANCE TEST AGAINST ZYMV (strains 13.18,FN and E15) on T_1 PROGENY FROM PLANTS GENETICALLY TRANSFORMED WITH pBPL4186

Name of the T_0 plant	Number of T_1 plants tested	Number of non-scorable plants	"FN reaction" or susceptible plants	Number of resistant plants	Most probable genetic determinism of resistance
Strain 13.18, FN					
Z4.1	40	4	36	0	/
Z5.1	38	0	38	0	/
Z5.2	27	16	11	0	/
Z12.7	40	0	32	8	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Z12.8	28	0	11	17	inoculation problem
Strain E15					
Z4.1	46	0	46	0	/
Z5.1	34	0	34	0	/
Z12.8	19	0	19	0	/

TABLE 6. RESULTS OF THE RESISTANCE TEST AGAINST ZYMV (strains 13.18,FN and E15) on T₁ PROGENY FROM PLANTS GENETICALLY TRANSFORMED WITH pBPL4194

Name of the T₀ plant	Number of T₁ plants tested	Number of non-scorable plants	"FN reaction" or susceptible plants	Number of resistant plants	Most probable genetic determinism of resistance
Strain 13.18, FN					
Z1.9	36	0	36	0	/
Z2.5	41	0	41	0	/
Z2.6	44	0	40	4	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Z2.7	47	4	43	0	/
Z2.23	46	0	40	6	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Z18.6	45	0	29	16	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Z22.9	20	0	15	5	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Strain E15					
Z1.9	40	0	40	0	/
Z2.5	41	0	41	0	/
Z2.6	9	0	9	0	/
Z2.7	38	0	38	0	/
Z22.9	19	0	19	0	/
Z22.13	12	8	4	0	/

TABLE 7. RESULTS OF THE RESISTANCE TEST AGAINST ZYMV (strains 13.18, FN and E15) on T$_1$ PROGENY FROM PLANTS GENETICALLY TRANSFORMED WITH pBPL4771

Name of the T$_0$ plant	Number of T$_1$ plants tested	Number of non-scorable plants	"FN reaction" or susceptible plants	Number of resistant plants	Most probable genetic determinism of resistance
Strain 13.18, FN					
Z23.15	28	5	15	8	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Z28.3	43	3	20	20	insertion at 1 locus of a resistance gene efficient in the heterozygous condition
Strain E15					
Z23.15	39	6	20	13	insertion at 1 locus of a resistance gene efficient in the homozygous condition
Z23.22	36	3	8	25	insertion at 1 locus of a resistance gene efficient in the heterozygous condition
Z25.1	12	1	10	1	
Z25.14	11	6	3	2	/
Z28.3	40	11	5	24	insertion at 1 locus of a resistance gene efficient in the heterozygous condition
Z31.2	22	1	18	3	/

TABLE 10

Observation data and ELISA values from transformed melon field trial (Season N° 2) (pBPL 3817)

VARIETY	N°BLOC	Notation n°1				Notation n°2				Notation n°3				Notation n°4				Notation n°5			
		Virus	W M V 2	C M V	Z Y M V	Virus	W M V 2	C M V	Z Y M V	Virus	W M V 2	C M V	Z Y M V	Virus	W M V 2	C M V	Z Y M V	Virus	W M V 2	C M V	Z Y M V
1	B1	0				3	1,36	-		6	1,20	-	-	6	2,41	2,13	-	NJ			-
1	B2	2				3				4				5				NJ			-
1	B3	2	0,85	-	-	3				4				5				NJ			-
2	B1	0				0				0				0	0,12	1,99	-	2(pc)		1,0-2,31	-
2	B2	0				0				0				2 plt 5	0,20	2,55	-	5		0,15	-
2	B3	0				0				0	0,06	1,49	-	2	0,11	1,87	-	NJ			-
3	B1	0				0				0				0				2(pc)		0,75-1,06	-
3	B2	0				0				0	0,09	1,34	-	2	0,18	2,59	-	1 plt 4			-
3	B3	0				0	0,02	0,08	-	0				0	0,17	1,56	-	NJ		0,23	-
4	B1	0				0				0				0	0,11	1,98	-	2(pc)		0,14	-
4	B2	0				1				0				0	0,09	0,95	-	2(pc)		0,16	-
4	B3	0				1				0				2 plt 3	0,05	0,81	-	NJ		1,7	-
5	B1	0				0				0				0				1(pc)		0,9	-
5	B2	0				0				0				0				2(pc)		0,18	-
5	B3	0				1	0,04	0,22	-	0				0				2(pc)		0,58-1,4	-
6	B1	2	0,93	0,76	-	2	0,98	-	-	4	1,08	-	-	5	1,62	0,97	-	6			-
6	B2	2	1,08	-	-	3				4				5				6			
6	B3	0				4				5	1,05	1,51	-	4				5			
Optic density value	Safe control X2 (2x)		0,04	0,12	0		0,04	0,12			0,04	0,65			0,13	0,67	0,02		0,45	0,1	
	Inoculated control		2,17	2,13			2,17	2,13			1,56	2,37			1,56	2,37	1,37		2,23	0,61	
	% of viral infection		40,5	5,4	0		97,3	27,7	8		100	100	0		100	100	0		100	100	0

The samples with optic density higher than the value of the safe control X2 were considered as positive

[shaded] : samples considered as negative

pc : chlorotic point

NJ : can not be judged

plt : plant

83

TABLE 9
Observation data and ELISA values
from transformed melon field trial (Season N° 1)
(pBPL 3817)

VARIETY	N°BLOC	Notation n° 1				Notation n° 2				Notation n° 3			
		Virus	WMV2	CMV	ZYMV	Virus	WMV2	CMV	ZYMV	Virus	WMV2	CMV	ZYMV
1	B1	0				0				0			
	B2	5	0,1	0,4	0	4	0,02	1,0	0	4	0,01	1,9	0
	B3	0				0				0			
2	B1	0				0				0			
	B2	0				1	0	0	0	3	0,01	1,4	0
	B3	1	0	0	0	4	0,03	0,9	0	0			
3	B1	0				0				0			
	B2	0				0				2	0,05	0,8	0
	B3	0				1	1.6???	0,7	0	1	0,02	0,7	0
4	B1	0				0				3	0,03	1,6	0
	B2	0				0				0			
	B3	0				5	0,09	1,1	. 0	0			
5	B1	0				0				0			
	B2	0				1	0,02	1,8		3	0,03	1,4	0
	B3	4	0	0,8	0	5	0,03	1,4		5	0,03	0,7	0
6	B1	0				0				0			
	B2	0				3	0,04	1,9	0	NJ	0,01	0,7	0
	B3	0				0				0			
7	B1	1	0	0	0	0				0			
	B2	0				2	0,03	1,7	0	NJ			
	B3	0				0				NJ			
8	B1	0				0				3	0,04	1,5	0
	B2	0		.		4	0,17	0,9	0	0			
9	B1	0				0				0			
	B2	0				0				0			
	B3	1	0	0	0	0				0			
10	B1	0				0				0	0,01	0,9	0
	B2	0				4	0,03	0,6	0	5	0,09	0,1	0
	B3	0				3	0,19	0,6	0	5			
11	B1	1	0	0	0	0				5	0,02	1,0	0
	B2	0				1	0,05	1,4	0	6	0,02	0,9	0
	B3	6	0	0,5	0	6	0,02	1	0	6	0,01	0,4	0
12	B1	5	1	0	0	4	2,1	0,15	0	0	1,86	0,15	0
	B2	1	2,1	0	0	3	2,2	0,17	0	2	2,16	1,67	0
	B3	5	2,2 ·	0	0	4	2,1	0,8	0	4	1,77	1,28	0
Optic density	Safe control X2 (x2)		0,03	0,02			0,05	0,05			0,02	0,15	
Value	Inoculated control		1,6	2,2			1,8	2,4			1,46	2,36	
	% of viral infection		92	24	0								

84

11. Transient assay of anti-viral ribozyme genes activity

11.1. Description of the system

The transient assay system for in vivo activity of ribozyme or antisense genes was described by Sawyer and Birch (1996). A target reporter gene and the ribozyme gene to be tested are cloned on two separate plasmids and co-introduced into Nicotiana suspension culture cells by particle bombardment. The target reporter gene is an in-frame fusion between the target sequence and the LUC sequence. This in frame fusion encodes a chimeric protein with luciferase activity. 48 hours after the plasmids delivery, the cells are placed under a camera which measures light emission; light emission represents the level of expression of the target reporter gene. The ribozyme gene tested expresses transiently a ribozyme which interacts with the target sequence and decreases the quantity of the chimeric target-LUC mRNA which is detected as a reduced luciferase activity.

To measure the activity of various ribozyme genes protecting transgenic melon plants against potyvirus infection, we have built various target reporter genes with viral sequences fused to LUC.

11.2. Genes used (see figure 15)

Target sites on the positive strand of the virus WMV2 (pBPL0188):

A 540 bases fragment containing four sites targeted by the polyribozyme (2259, 2403, 2520, 2720) was amplified by PCR and cloned in-frame in front of LUC, under control of a modified version of the 35S promoter, using the plasmid pGLFuz (see figure 17) as described by Sawyer and

Birch (1996). Because the most 3' site is located close to the stop codon, it was also necessary to change one base of the viral sequence to remove this stop codon and provide a flanking sequence to this site.

Target sites on the negative strand of the virus WMV2 (pBPL0184 to 0187):

Because the negative strand is not coding, it was impossible to clone a large sequence containing several ribozyme sites in frame with LUC. Four shorter sequences (between 180 and 220 bases in length) each containing one ribozyme site (696, 996, 1747, 2067) were amplified by PCR and cloned in-frame in front of LUC, under control of the modified version of the 35S promoter in pGLFuz as above.

Ribozyme gene against WMV2 (pBPL0189):

The short arm polyribozyme targeting four sites on the plus (2259, 2403, 2520 and 2720) and four sites on the minus strand (696, 996, 1747 and 2067) of WMV2 from pBPL4743 was cloned under control of the modified version of the 35S promoter in pGLFuz.

Ribozyme gene against ZYMV (pBPL0190):

One site on the minus strand (hereafter referred to as 1747) of the virus is present on both WMV2 and ZYMV. The pentameric ribozyme against the minus strand of ZYMV in pBPL4747 is targeting site 1747 amongst others. In this sequence, the ribozyme against site 1747 is identical to the ribozyme present in the construct against WMV2, but the arms are shorter: 8 bases each side. The pentameric ribozyme was cloned under control of the modified version of the 35S promoter in pGLFuz.

11.3. Assay and results (see figure 16)

86

The assays were carried out as described by Sawyer and Birch (1996). For each assay, target and ribozyme plasmid DNA (2μg each, approx. 25% molar excess of ribozyme) were co-precipitated onto tungsten particles and shot onto Nicotiana plumagnifolia cells on filter paper discs (harvested from suspension culture). After bombardment, the cells were cultured on the discs for 48 hours at 28° C in the dark. For luciferase assay, the cells were placed into a petri dish containing D-luciferin for 30 minutes, then imaged using a liquid nitrogen cooled CCD camera and AT1 software (Wright Instruments Ltd, Enfield, UK). Results are shown in figure 16.

The WMV2 polyribozyme (pBPL0189) could inhibit the target containing the four sites on the plus strand (pBPL0188) to an extent (70%) not seen in previous uses of the system with full length antisense when an equal amount of target and inhibitor gene was used (Sawyer and Birch, 1996).

Both ribozymes (pBPL0189 and pBPL0190) could inhibit the target containing site 1747 (pBPL0186) to an even greater extent (75%). Site 696 (pBPL0184) seemed not amenable to inhibition; this absence of effect on one target shows that the effect observed is not simply caused by an interaction between the 2 plasmids or their promoter. Sites 996 (pBPL0185) and 2067 (pBPL0187) were slightly inhibited.

11.4. Comparison with in vitro data obtained previously

Minus strand target (see figure 12):

Site 696 was the site least cleaved in vitro by the WMV2 ribozyme, and was not cleaved at all at 22°C. This is in agreement with the transient assay results.

Sites 1747 and 996 seemed equally cleaved in vitro and 2067 was the site most amenable to cleavage in vitro,

87

particularly at 22°C. This is not fully consistent with
the transient assay results suggesting that 1747 is
clearly more amenable to cleavage than the 2 others. The
comparison of the in vitro and transient in vivo results
suggests that the in vitro activity is not a perfect
predictor of the transient activity in vivo (unable to
rank sites 996, 1747, 2067), but it has some predictive
capacity (it could predict that 696 would not be a good
target).

Plus strand target (see figure 13):

When the ribozyme was tried in vitro on a target of
about 900 bases, site 2259 was somewhat cleaved at 22°C
and 37°C (with 4.5 - 13 fold excess ribozyme) and to a
lesser extent site 2403; cleavage was far from complete.
At 50°C all four sites were cleaved, 2259 then 2403 being
apparently the fastest. The transient in vivo assay shows
that this ribozyme is active despite the low level of
activity observed in vitro at low temperature.
ZYMV Ribozyme:

When the ZYMV ribozyme was assayed in vitro on a
ZYMV target, site 1747 was not cleaved at 22°C. This
result cannot be compared directly with the in vitro assay
on WMV2 since the target sequence is significantly
different, except immediately around the site itself; it
suggests however that shorter arms reduce the activity in
vitro. Results of the transient assay suggest that at
least for this particular site 1747, the arm length can be
reduced with no significant effect on activity in vivo.
The predictive capacity of the in vitro assay for this
system is therefore not good.

11.5. Construction of pGLFuz (see Figure 17):

To create a uidA::luc translational fusion vector
(pGLFuz) the uidA gene was first modified by PCR

88

amplification to introduce novel XhoI and NsiI sites
upstream of the start codon, replace the stop codon with
an alanine codon (TGA -> GCG) and introduce a novel NotI
site downstream of this codon. The modified uidA gene PCR
product was ligated into pGemT (Promega). The appropriate
recombinant plasmid was designated pGemGUS.

The modified uidA gene from pGemGUS was ligated into
pZorz vector as a NsiI - NcoI fragment. In the resulting
plasmid clone, pGLFuzTAG, the uidA coding region is
flanked by an osa promoter and ocs polyadenylation region.

Due to the presence of a stop codon between the uidA
and luc genes, preventing a translational fusion, the luc
gene in pGLFuzTAG was removed (by digestion with NotI,
treatment with T4 DNA polymerase to create blunt ends,
then digestion with XhoI) and replaced by the luc gene
from pSPNF+luc (Promega) (prepared by digestion with KpnI,
treatment with T4 DNA polymerase and digestion with XhoI).
The appropriate recombinant plasmid was designated pGLFuz.

For construction of target genes, the uidA coding
sequence is replaced by viral sequences. For construction
of ribozyme genes, both the uidA and luc coding sequences
are replaced by the ribozyme sequences. (see figure 15)

REFERENCES

Betzner et al., 1997, The Plant Journal, vol. 11, n° 3,
p. 101-109.

Bos L (1969). Experience with a collection of plant
viruses in leaf material dried and stored over calcium
chloride and a discussion of litterature on virus

89

preservation. Medec. Fac. Landbouwwet. Rijksuniv. Gent. 34:875-887

Firek et al., 1993, A wound-induced promoter driving nptII expression limited to dedifferentiated cells at wound sites is sufficient to allow selection of transgenic shoots. Plant Mol. Biol., vol. 22, p 129-142.

Fromm M.E., Taylor L.P. and Walbot V., 1986. Nature, vol. 319, p. 791.

Guerineau F. and Mullineaux P., 1993, Chapter 4 "Plant transformation and expression vectors" in Plant Molecular Biology LABFAX, ed. Croy R.R.D., BIOS Scientific Publishers, Blackwell Scientific Publications.

Haseloff and Gerlach, 1988, Simple RNA enzymes with new and highly specific endoribonuclease activities. Nature, vol. 334, p. 585-591.

Lecocq H and Pitrat M (1984). Strains of ZYMV in Muskmelon (C. melo L.). Phytopath. Z., 11:165-173

Marrov J (1967). Amélioration des méthodes de transmission mécanique des virus par absorption des inhibiteurs d'infection sur le charbon végétal. C. R. Acad. Agric. Fr. 53:972-981

Pascual Perez, Gérard Tiraby, Jean Kallerhoff and Joel Perret, 1989, Phleomycin resistance as a dominant selectable marker for plant cell transformation. Plant Molecular Biology, vol. 13, p 365-373.

Perriman et al, 1992, Extended target-site specificity
for a hammerhead ribozyme. Gene, vol. 113, p. 157-163.

Sawyer B. and Birch R., 1996, A generic transient assay
for antisense and ribozyme inhibition in vivo, vol 28,
1996, Proceedings of the Australian Society for
Biochemistry and Molecular Biology.

91

CLAIMS

1. Nucleic acid molecule, called polyribozyme, having endoribonuclease activity, comprising at least first and second hybridising regions and at least two catalytic regions, said first and second hybridizing regions being complementary to first and second target regions, respectively, in the genomic (+) RNA of a potyvirus, or in the replicative (-) strand RNA of a potyvirus, said first and second target regions being either :
- within or comprising different genes or untranslated regions on the same RNA strand, or
- on different RNA strands, or
- within a region corresponding to one gene or cistron on the replicative (-) strand RNA.

2. Nucleic acid molecule called polyribozyme, having endoribonuclease activity, comprising at least two units « R_1 » and « R_2 », each of which has the general formula :

$$3' \ [(X)_n - Q - (X')_{n'}]_p \ 5'$$

wherein :
- X and X' each represent, independently, any ribonucleotide ; $(X)_n$ and $(X')_{n'}$ represent ribonucleotide sequences substantially complementary to at least part of a target region of potyvirus genomic (+) RNA or to at least part of its (-) strand replicative RNA ;
- the complementary sequences $(X)_n$ and $(X')_{n'}$ of R_1 are complementary to sections of a first target region, Z_1, of the potyvirus genomic (+) RNA or of the replicative (-) strand RNA, region Z_1 being within or comprising a first gene or untranslated region (UTR) ;

92

and $(X)_n$ and $(X')_{n'}$ of R_2 are complementary to sections of a second target region, Z_2, of the potyvirus genomic RNA or (-) strand replicative intermediate, region Z_2 being within or comprising a second gene or untranslated region ;

Z_1 and Z_2 being either :

- different genes or untranslated regions, on the same RNA strand, or
- being the same or different genes or untranslated regions on different strands of the potyvirus RNA, or
- being the same genes on the (-) strand replicative RNA ;
- n and n' are each equal to or greater than 4, preferably equal to or greater than 8 ;
- Q represents the catalytic region of a ribozyme ;
- $p \geq 2$.

3. Nucleic acid molecule according to claim 2 wherein each of the R units is represented by the following formula :

```
    3'                                          5'
   (X)n  Y                          Y  (X)n'
         A                          C
         A                          U
         G          1A              G
         C  *   G        G          A
         |       |          X
         Y'      Y''
```
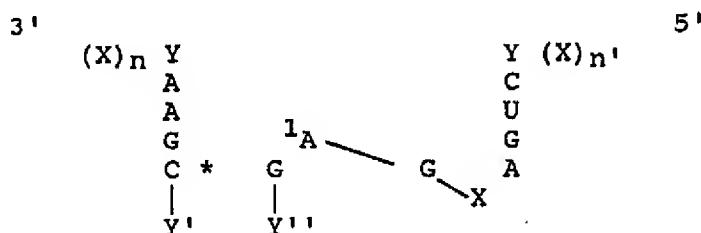
wherein, each X represents a ribonucleotide which may be the same or different, $(X)_n$ and $(X)_{n'}$ represent ribonucleotide sequences each substantially complementary to one of two RNA sequences in a potyvirus genome or replicative (-) strand RNA, the said two potyvirus sequences being separated from each other by at least one nucleotide in the genome, or in the replicative (-) strand RNA; an (*) represents a base pair between complementary ribonucleotides ; Y represents any ribonucleotide and each Y may be the

same or different ; Y' and Y" represent ribonucleotides
of complementary sequence along at least part of their
length to allow base pairing between the
oligoribonucleotides, or Y' and Y" together form a
single RNA sequence wherein at least part of said
sequence comprises a stem formed by base pairing
between complementary nucleotides ; and optionally, an
additional nucleotide selected from any one of A, G, C
or U is inserted after ^{1}A.

4. Nucleic acid molecule according to claim 2 or
3 wherein the complementary sequences $(X)_n$ and $(X')_{n'}$
of R_1 are complementary to a region within or
comprising a first gene or U.T.R. on the genomic (+)
RNA, and X and X' of R_2 are complementary to a region
within, or comprising, a second gene or U.T.R. on the
genomic (+) RNA.

5. Nucleic acid molecule according to claim 2 or
3 wherein the complementary sequences $(X)_n$ and $(X')_{n'}$
of R_1 are complementary to a region within or
comprising a region corresponding to a first gene OR
U.T.R. on the replicative (-) strand RNA, and X and X'
of R_2 are complementary to a region within or
comprising a region corresponding to a second gene or
U.T.R. on the replicative (-) strand RNA.

6. Nucleic acid molecule according to claim 2 or
3 wherein the complementary sequences $(X)_n$ and $(X')_{n'}$
of R_1 are complementary to a region Z_1 within or
comprising a gene or U.T.R. on the genomic (+) RNA, and
X and X' of R_2 are complementary to a region Z_2 within
or comprising a region corresponding to a gene or
U.T.R. on the replicative (-) strand RNA.

7. Nucleic acid molecule according to claim 6
wherein the complementary sequences of R_1 are
complementary to a region Z_1 within or comprising a
cistron « C » on the (+) strand RNA and the
complementary sequences of R_2 are complementary to a

94

region Z_2 on the (-) strand within or comprising the region corresponding to the same cistron « C ».

8. Nucleic acid molecule according to claim 6 comprising further R units whose complementary sequences $(X)_n$ and $(X')_{n'}$ are complementary to regions within, or comprising, further genes on the genomic (+) strand or on the replicative (-) strand, and/or to regions within the same genes as Z_1 and/or Z_2.

9. Nucleic acid molecule according to any one of the preceding claims wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of at least one R unit are complementary to regions within or comprising the capsid protein gene, the NIa gene, the NIb gene, the 3' untranslated region of a potyvirus, or the corresponding sequences in the replicative (-) strand RNA.

10. Nucleic acid molecule according to any one of claims 1 or 2 to 9, wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of at least one R unit are complementary to regions within or comprising the capsid protein (CP) gene, or the corresponding sequences in the replicative (-) strand RNA.

11. Nucleic acid molecule according to claim 10, wherein the complementary segments $(X)_n$ and $(X')_{n'}$ of R_1 are complementary to regions within or comprising the CP gene on the replicative (-) strand RNA.

12. Nucleic acid molecule according to claim 10, wherein the complementary segments $(X)_n$ and $(X')_{n'}$ of R_1 are complementary to regions within or comprising the CP gene on the genomic (+) strand RNA.

13. Nucleic acid molecule according to claim 11 wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of both R_1 and R_2 are complementary to regions within the CP gene on the replicative (-) strand RNA.

14. Nucleic acid molecule according to claims 4, 10 and 12 wherein the complementary sequences $(X)_n$ and

$(X')_{n'}$ of R_1 and of R_2 are respectively complementary
to regions within the CP gene and the 3' non-translated
region of the genomic (+) strand RNA.

15. Nucleic acid molecule according to claims 5
and 9 wherein the complementary sequences $(X)_n$ and
$(X')_{n'}$ of R_1 and R_2 are respectively complementary to
regions of the replicative (-) strand RNA corresponding
to regions within the CP gene and the NIb gene, or to
the NIb gene and the NIa gene.

16. Nucleic acid molecule according to claims 11
and 12 comprising at least two R units, R_1 and R_2,
wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of
R_1 are complementary to the CP gene of the genomic (+)
strand RNA, and the hybridising regions of R_2 are
complementary to the CP gene on the replicative (-)
strand RNA.

17. Nucleic acid molecule according to claims 8
and 9 comprising at least three R units, R_1, R_2 and R_3,
wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of
R_1 are complementary to regions within the CP gene or
the 3' non-translated region (3'NT) of the genomic (+)
strand RNA, the complementary sequences $(X)_n$ and $(X')_{n'}$
of R_2 are complementary to regions within the NIb gene
on the replicative (-) strand RNA and the complementary
sequences $(X)_n$ and $(X')_{n'}$ of R_3 are complementary to
regions within the NIa or CP genes on the replicative
(-) strand RNA.

18. Nucleic acid molecule according to claim 2 or
3, comprising at least two R units, « R_1 » and « R_2 »,
wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of
R_1 are complementary to sections of a target region Z_1
within a region corresponding to a given gene or
untranslated region in the replicative (-) strand RNA,
and the hybridising sequences X and X' of R_2 are
complementary to a target region Z_2 within the same

96

gene or untranslated region as that of Z_1, on the same RNA strand.

19. Nucleic acid molecule according to any one of claims 1 or 2 to 18 wherein « p » has a value from 2 to 20, for example 4 to 10.

20. Nucleic acid molecule according to any one of claims 1 or 2 to 19 wherein each of the complementary sequences $(X)_n$ and $(X')_{n'}$ of at least one R unit comprise, independently, greater than 50 nucleotides, for example 60 to 300 nucleotides.

21. Nucleic acid molecule according to any one of claims 1 or 2 to 19 wherein each of the complementary sequences $(X)_n$ and $(X')_{n'}$ of at least one R unit comprise, independently, from 4 to 50 nucleotides, for example from 8 to 12 nucleotides.

22. Nucleic acid molecule according to claims 20 and 21 wherein at least a first R unit has complementary sequences $(X)_n$ and $(X')_{n'}$ each comprising more than 50 nucleotides, and at least a second R unit has complementary sequences $(X)_n$ and $(X')_{n'}$ each comprising from 4 to 50 nucleotides.

23. Nucleic acid molecule according to any one of claims 1 or 2 to 22 wherein the 3' $(X)_n$ sequence of a given R unit and the 5' $(X')_{n'}$ sequence of the adjacent R unit are complementary to sequences which are contiguous in the potyvirus genomic (+) strand RNA or replicative (-) strand RNA.

24. Nucleic acid molecule according to any one of claims 1 or 2 to 22 wherein the 3' $(X)_n$ sequence of a given R unit and the 5' $(X')_{n'}$ sequence of the adjacent R unit are complementary to non-contiguous portions of the potyvirus genomic (+) strand RNA or replicative (-) strand RNA.

25. Nucleic acid molecule according to any one of claims 1 or 2 to 24 wherein at least one nucleotide, or a sequence of nucleotides « A », non-complementary to

97

the viral genome or the replicative (-) strand RNA, is included between 2 consecutive R units.

26. Nucleic acid molecule according to claim 25 wherein the non-complementary sequence « A » is a polylinker or a spacer.

27. Nucleic acid molecule according to any one of claims 1 or 2 to 26 wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ of a first R unit and the complementary sequences $(X)_n$ and $(X')_{n'}$ of a second R unit are complementary to target sequences in the genome or in the (-) sense RNA of at least 2 distinct species of potyvirus.

28. Nucleic acid molecule according to claim 1 or 2 wherein one or more catalytic region comprises, or is derived from a hairpin ribozyme catalytic region.

29. Nucleic acid molecule according to any one of claims 1 to 27 wherein the hybridising regions are complementary to target regions of one or more potyviruses chosen from the aphid-, mite- and whitefly transmitted potyviruses including

Zucchini yellow mosaic virus	(ZYMV),
Watermelon mosaic virus 2	(WMV2),
Tobacco vein mottling virus	(TVMV),
Tobacco etch virus	(TEV),
Potato virus Y	(PVY),
Pepper mottle virus	(PepMoV),
Soyabean mosaic virus	(SbMV),
Papaya ringspot virus	(PRSV),
Pea seed-borne mosaic virus	(PSbMV),
Turnip mosaic virus	(TuMV),
Johnson grass mosaic virus	(JGMV),
Plum pox virus	(PPV)
Maize Dwarf Virus	(MDV)
Bean common mosaic	(BCMV)
Bean yellow mosaic	(BYMV)
Carnation vein mottle	(CarVMV)

98

Celery mosaic	(CeMV)
Clover yellow vein	(ClYVV)
Henbane mosaic	(HMV)
Lettuce mosaic	(LtMV)
Sugarcane mosaic	(SCMV)
Peanut stripe virus	(PStV)

30. Nucleic acid molecule according to claim 29 wherein the complementary sequences $(X)_n$ and $(X')_{n'}$ are complementary to target regions of ZYMV and/or WMV2, or to mutants or derivatives thereof.

31. Nucleic acid molecule according to any one of claims 1 to 30 further comprising at its 3' end, all or part of the 3' untranslated region of a potyvirus, particularly ZYMV or WMV2.

32. Nucleic acid molecule according to any one of claims 1 to 31, capable of cleaving potyvirus genomic (+) strand RNA or replicative (-) strand RNA thereby reducing viral replication, infection, assembly, cell to cell movement and/or symptoms.

33. Nucleic acid molecule consisting of nucleotides 1704 to 1873 of the (-) strand of the sequence illustrated in Figure 18, or a fragment thereof consisting of at least 20 consecutive nucleotides of the said molecule, said fragment including the GTA shown at position (-) 1748 to (-) 1750.

34. Nucleic acid molecule consisting of a sequence having at least 75% identity and preferably at least 85% identity with the molecule according to claim 33, including the GTA shown at position (-) 1748 to (-) 1750 of Figure 18.

35. Nucleic acid molecule comprising any of the molecules according to claim 33 or 34, but which is not an entire viral (-) sense replicative RNA.

36. Nucleic acid molecule, comprising a sequence complementary to the molecules of claim 33 or 34.

37. Nucleic acid molecule comprising the molecule of claim 33 or 34 linked as a translational or transcriptional fusion to a second nucleic acid molecule.

38. Nucleic acid molecule according to claim 37 wherein the second nucleic acid molecule comprises a coding sequence, encoding for example a reporter protein, or any other gene of interest.

39. Nucleic acid molecule according to any one of claims 1 to 3 having endoribonuclease activity, wherein said molecule is capable of hybridising to, and cleaving, a molecule according to claims 33-35 or 37-38.

40. Chimeric gene comprising a DNA sequence which is operably linked to 5' and 3' transcription regulation signals functional in plant cells and which, upon transcription, gives rise to one or more nucleic acid molecule(s) according to any one of claims 1 to 39.

41. Vector comprising a nucleic acid molecule, or which when transcribed gives rise to a nucleic acid molecule according to any one of claims 1 to 39, or a chimeric gene according to claim 40, and optionally a selectable marker gene, reporter gene, and/or functional sequences capable of transferring DNA to a plant cell genome.

42. Plant cells comprising a nucleic acid molecule according to any one of claims 1 to 39, or a chimeric gene according to claim 40, or a DNA sequence which, upon transcription, gives rise to a nucleic acid molecule according to claims 1 to 39.

43. Transgenic plant comprising a nucleic acid according to any one of claims 1 to 32 or a chimeric gene according to claim 40, and capable of reducing replication, infection, cell-to-cell movement and/or assembly of a potyvirus.

44. Transgenic plant comprising a nucleic acid according to any one of claims 33 to 39.

45. Transgenic plant according to claim 43, wherein the plant is chosen from the following generae : Cucurbitaceae, Chenopodiaceae, Cruciferaceae, Graminaceae, Leguminoseae, Ombelliferaceae, Rosaceae, Solanaceae

46. Transgenic plant according to claim 45, said plant being courgette, melon, beet, oat, wheat, sugar cane, barley, ryegrass, bean, pea, soya, peanut, carrot, celery, parsnip, plum, potato, tomato, tobacco, pepper.

47. Seeds of transgenic plants according to any one of claims 43 to 46.

48. Progeny of transgenic plants according to any one of claims 43 to 46.

49. Fruit of transgenic plants according to any one of claims 43 to 46.

50. Process for conferring resistance to a potyvirus on a plant, comprising introducing into the plant, or into a part of the plant, a nucleic acid comprising, or encoding, a nucleic acid according to any one of claims 1 to 32, optionally followed by regeneration of a transgenic plant exhibiting the required resistance, and optionally reproduction to produce progeny.

51. Method for controlling the functional expression of a nucleic acid molecule by

i) translationally or transcriptionally fusing an artificial target sequence comprising a nucleic acid according to any one of claims 33 to 35, to the nucleic acid whose expression is to be controlled, giving rise to a hybrid gene whose expression and function are not disrupted by the fusion,

ii) allowing the transcription product of the hybrid gene to come into contact with an RNA molecule

complementary to at least a part of the artificial
target sequence, and capable of hybridising to, and/or
cleaving the artificial target.

52. Method according to claim 51 wherein the
complementary RNA molecule is a ribozyme or antisense
molecule.

53. Nucleic acid molecule, called « ribozyme »,
having endoribonuclease activity, comprising at least
one hybridising region and at least one catalytic
region, said hybridising region being complementary to
one or more target regions of the genomic (+) RNA of a
potyvirus, and/or to one or more target regions of the
replicative (-) strand RNA of a potyvirus, said target
region consisting or comprising the nucleic acid
molecule of claim 33 or 34.

54. Nucleic acid molecule according to claim 53
having one catalytic region.

FIGURE 1

Amplification of potyviral genes
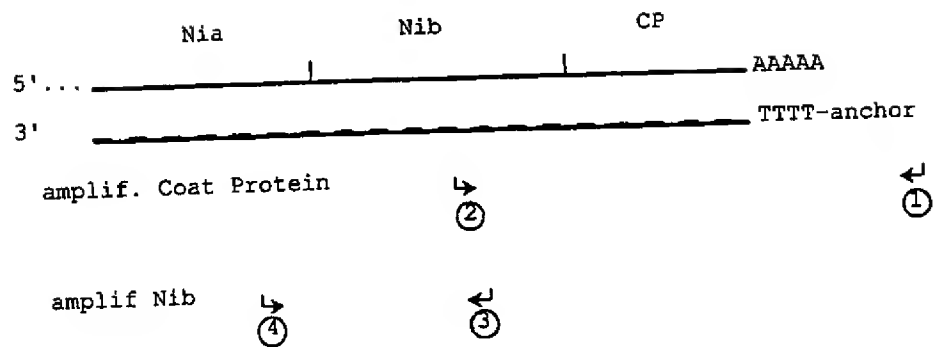
Primers: 1: anchor

          2: coding, TVVDNTLMV

          3: reverse, ANGDDIIL

          4: coding, FWIHWITT

Plain line: viral RNA, 3cm/kb approx.

Dashed line: cDNA

FIGURE 2 (a)

WMV2 USA: Quemada et al, J. Gen. Virol. 1990, 71, 1451-1460.
G   N   N   S   G   Q   P   S   T   V   V   D   N   T   L   M   V   V   I A
AGGGAACAACAGTGGGCAGCCATCTACAGTTGTTGACAATACACTTATGGTT
(1577) (base number in partial published sequence)

ZYMV: Quemada et al, J. Gen. Virol. 1990, 71, 1451-1460.
G   N   N   S   G   Q   P   S   T   V   V   D   N   T   L   M   V   V   I   S
GGTAACAACAGTGGGCAACCCTCAACAGTGGTGGATAACACACTAATGGTT
(260) (as above)

Plum Pox Virus NAT: Maiss et al, J. Gen. Virol. 1989, 70,
513-524.
2599 (amino acid number from the complete published sequence)
G   N   N   S   G   Q   P   S   T   V   V   D   N   T   L   M   V   I   L   A
GGAAACAATAGTGGTCAACCCTCGACAGTTGTTGACAATACACTCATGGTT
7941 (base number from the complete published sequence)

TVMV: Domier et al, N.A.R., 1986, 14, 13, 5417-5429.
2531 (as above)
G   N   N   S   G   Q   P   S   T   V   V   D   N   T   L   M   V   V   L   A
GGCAACAACAGTGGACAGCCATCTACGGTTGTGGACAACACACTCATGGTG
7796 (as above)

TEV: Allison et al, Virology, 1986, 154, 9-20.
2586 (as above)
G   N   N   S   G   Q   P   S   T   V   V   D   N   T   L   M   V   I   I   A .
GGCAACAATAGCGGGCAACCTTCAACAGTGGTGGACAACACACTCATGGTC
7900 (as above)

PVY: Robaglia et al, J. Gen. Virol. 1989, 70, 935-947.
2584 (as above)
G   N   N   S   G   Q   P   S   T   V   V   D   N   S   L   M   V   V   L   A
GGTAATAATAGTGGTCAACCTTCTACCGTTGTGGATAATTCTCTCATGGTT
7934 (as above)

FIGURE 2 (b)

PRSV: Yeh et al, J. Gen. Virol. 1992, 73, 2531-2541.
2826 (as above)
G   N   N   S   G   Q   P   S   T   V   V   D   N   T   L   M   V   L   I   T
GGCAATAATAGTGGCCAACCTTCGACAGTTGTTGATAATACATTGATGGTTTTAATC
8561 (as above)

PSbMV: Johansen et al, J. Gen. Virol.1991, 72, 2625-2632.
2702 (as above)
G   N   N   S   G   Q   P   S   T   V   V   D   N   T   L   M   V   I   L   A
GGAAATAATAGTGGACAACCATCAACAGTTGTTGATAACACATTGATGGTGATTCTT
8247 (as above)

TuMV: Nicolas et al, J. Gen. Virol.1992, 73, 2785-2793
2665 (as above)
G   N   N   S   G   Q   P   S   T   V   V   D   N   T   L   M   V   I   L   A
GGAAACAATAGCGGACAGCCATCGACTGTTGTAGACAACACACTCATGGTCATATTG
8123 (as above)

FIGURE 3 (a)

WMV2 USA
F F A N G D D I I L + 161aa to CP; then ca. 1100 bp to end
TTCTTTGCCAATGGTGATGATATTATCCTG
(1697)

ZYMV
F F A N G D D L I L + 161aa to CP; then ca 1050 bp to end
TTCTTTGCAAATGGAGATGATCTGATACTT
(381)

Plum Pox Virus NAT
2639
Y F V N G D D L V L + ca. 1670 bp to the end, 162 aa to CP
TACTTTGTGAATGGTGATGATCTTGTCCTC
8061

TVMV
2571
F F A N G D D L I I + ca 1550 bp to the end, 160 aa to CP
TTCTTTGCCAATGGAGATGATCTTATAATT
7916

TEV
2622
Y Y V N G D D L L I + ca. 1450 bp to the end, 160 aa to CP
TATTACGTCAATGGCGATGACCTATTGATT
8008

FIGURE 3 (b)

PVY
2624
F F V N G D D L L I + ca.1640 bp to the end, 163 aa to CP
TTCTTTGTTAATGGTGATGACTTATTGATT
8054

PRSV
2866
F Y I N G D D L C I + ca. 1630 bp to the end, 162 aa to CP
TTTTATATCAATGGTGATGATCTCTGTATT
8681

PSbMV
2742
Y F A N G D D L L I + ca. 1530 bp to the end, 162 aa to CP
TATTTTGCAAATGGTGATGATCTATTGATA
8367

TuMV
2705
F F V N G D D L L L + ca. 1560 bp to the end, 161 aa to CP
TTCTTCGTCAACGGAGACGATTTACTGCTA
8242

FIGURE 4 (a)

WMV2 USA

                                              + 97 aa to the end

G   S   F   W   I   H   W   I   T   T   Q   D   G   F
        TTCTGGATACATTGGATAACGACT
        (344)

Plum Pox Virus NAT
        2188   + 93 aa to the end
S   H   F   W   K   H   W   I   S   T   K   D   G   H
        TTCTGGAAACACTGGATTAGTACG
        6711

TVMV
        2123   + 91 aa to the end
T   S   F   W   Q   H   W   I   T   T   K   D   G   Q
        TTCTGGCAACACTGGATAACAACA
        6572

TEV
        2176   + 92 aa to the end
G   I   F   W   K   H   W   I   Q   T   K   D   G   Q
        TTCTGGAAGCATTGGATTCAAACC
        6670

PVY
        2419   + 94 aa to the end
S   T   F   W   K   H   W   I   E   T   D   N   G   H
        TTCTGGAAGCATTGGATTGAAACA
        6692

PRSV
        2419   + 90 aa to the end
G   D   F   G   C   H   W   I   S   T   N   D   G   D

FIGURE 4 (b)

PSbMV
```
     2288                              + 96 aa to the end
G  K  Y  W  K  H  W  I  T  T  K  E  G  H
```


TuMV
```
     2254                              + 93 aa to the end
S  Q  F  W  K  H  W  I  S  T  K  D  G  Q
```

FIGURE 5 (a)

```
  1   GAATTCTAGA TTCTGGAAAC ATTGGATAAC CACTCAAGAT GGTTTTGTG GGCTACCTCT
 61   CGTGTCTGTC AATGATGGAT ATATTGTTGG TATACATGGA TTAACATCTA ATGATTCAGA
121   GAAGAATTTC TTTGTTCCAT TTACTGACGG ATTTGAAACG GAATACTTGA ACAATGCTGA
181   TAATCTGTCG TGGGACAAAC ACTGGTTTTG GGAACCAAGT AAGATCGCAT GGGGCTCTTT
241   GAATTAGTTA GAAGAACAGC CAAAGGAAGA ATTCAAGATA TCAAAACTCG TGTCTGATCT
301   TTTTGGCAAC ACGGTAGCAG TACAGAGTAG AAAAGAGAGA TGGGTTTTGG ATGCCATGGA
361   AGGGAATTTG ATAGCTTGTG GGCAAGCTGA GTGACAAAGC ACGTGGTTAA
421   AGGAAAAATGC CCTCATTTTG CACAATATCT TTCGTTACAT GATGAAGCAA AACAGTTCTT
481   TGAGCCATTG ATGGGGGCAT ACCAGCCAAG TCGGTTGAAT AAAGACGCTT TCAAGAAAGA
541   TTTCTTTAAG TACAACAAAC CGGTTGTTCT GAATGAAGTT GATTTTAATG CTTTCGAGAA
601   GGCAGTCGAG GGAGTGATAA CAATGATAGT TTTGCAGAAT GCTTATTTGT
661   GACTGATCCT GATGAGATTT ATGGATCCCT GAACATGAAA GCTGCTGTCG GTGCTCAATA
721   CAAAGGAAAG AAACAGGATT ACTTTTCTGG CATGGATAGT TTTGATAAGG AGCGCCTGTT
```

FIGURE 5 (b)

```
 781  GTACCTTAGC TGTGAGAGGC TTTTTAATGG AGAGAAAGGA ATTTGGAACG GTTCCTTGAA

 841  AGCGGAGTTG AGGCCAATTG AAAAAGTGCA GGCAAACAAA ACTAGAACAT TTACAGCAGC

 901  ACCGATTGAT ACATTGCTTG GAGCTAAAGT TTGCGTCGAT GATTTTAATA ATCAATTCTA

 961  CAGCTTCAAC TTAAAATGTC CATGGACAGT TGGCATGACT AAGTTTTATG GGGGTTGGGA

1021  TAAGCTAATG AGGAGTTTGC CTGATGGCTG GACATATTGC CACGCAGACG GTTCACAGTT

1081  CGACAGTTCT TTGACTCCCT TATTGTTGAA TGCTGTCCTC AGTATTAGAT GTTGCTTCAT

1141  GGAAGATTGG TGGGTTGGAA AGGAAATGCT TGAAAACCTT TACGCTGAGA TAGTTTACAC

1201  ACCAATCTTA GCACCTGATG GCACAATTTT CAAAAAGTTT AGAGGGAATA ACAGTGGACA

1261  GCCATCTACA GTTGTTGACA ATACGCTTAT GGTTGTCATT GCTATGTATT ACTCGTGCTG

1321  TAAACAAGGT TGGTCAGAGG AGGATATTGA AGAAAGGTTG GTGTTCTTTG CTAATGGTGA

1381  TGATATTATC TTGGCAGTTA GGGATGAGGA TGTGTGGCTG TATGACACTC TAAGTGCTTC

1441  ATTCGCTGAG CTTGGCCTGA ATTATAACTT CGATGAGCGG ACAAAGAAAA GAGAGGAGCT

1501  ATGGTTCATG TCACACCAAG CCATGCTAGT TGATGGGATT TACATTCCAA AGCTTGAACC
```

FIGURE 5 (c)

```
1561 TGAAAGGATA GTTCCATCT  TAGAGTGGGA  TAGGAGCAAG  GAACTTACGC  ATCGCACTGA
1621 GGCAATTGT  GCAGCCATGA  TTGAGGCATG  GGGACATACT  GAGCTTTTGC  AGGAAATTCG
1681 CAAATTTTAT TTGTGGCTTC  TAAGCAAAGA  TGAATTTAAG  GAACTCGCTG  CATCCGGAAA
1741 AGCTCCATAC ATAGCTGAGA  CAGCATTGAG  AAAGCTCTAC  ACAGATGTTA  ACACACAACC
1801 AAGTGAATTG CAAAAATACC  TTGAAGTTTT  GGATTTTAAC  CACATTGATG  GTTGCTGTGA
1861 ATCAGTGTCT CTGCAATCAG  GAAAAGAAGC  AGTAGAAAAT  TTGGATGCAG  GGAAGGACTC
1921 GAAGAAGGAC ACCAGTGGCA  AAGGGGATAA  GCCACAAAAT  TCGCAAACTG  GTCAAGGTAG
1981 CAAAGAACAG ACAAAAACTG  GTACAGTCAG  CAAAGATGTG  AATGTTGGAT  CGAAAGGAAA
2041 AGAAGTCCCA CGACTACAAA  AGATAACAAA  GAAAATGAAC  CTTCCAACAG  TTGGTGGGAA
2101 AATCATTCTT AGCTTAGACC  ATTTACTCGA  GTACAAACCT  AATCAAGTTG  ATCTGTTTAA
2161 CACTCGAGCA ACAAAAACAC  AGTTTGAATC  ATGGTACAGC  GCAGTTAAAG  TTGAATACGA
2221 TCTTAACGAT GAGCAAATGG  GTGTGATTAT  GAATGGTTTT  ATGGTTTGGT  GTATCGATAA
2281 CGGTACACTCT CCAGATGTCA  ATGGAGTTTG  GGTGATGATG  GATGGGGAAG  AGCAAGTTGA
```

FIGURE 5 (d)

2341 GTATCCATTA AAGCCAATTG TTGAAAATGC GAAACCAACT TTAAGACAAA TCATGCACCA

2401 TTTCTCAGAC GCAGCAGAAG CATATATTGA AATGAGAAAC TCTGAAAGTC CGTATATGCC

2461 TAGATACGGA TTACTAAGAA ATTTGAGAGA CAGGGAATTA GCACGCTATG CTTTTGACTT

2521 TTATGAGGTT ACTTCCAAAG CACCTAATAG GGCAAGAGAA GCAATAGCAC AAATGAAGGC

2581 CGCAGCTCTC GCGGGAATTA ACAGCAGGTT ATTTGGGCTT GATGGTAATA TCTCGACCAA

2641 TTCCGAAAAT ACTGAGAGGC ACACTGCAAG GGACGTGAAT CAGAATATGC ATACTTTGTT

2701 GGGTATGGGT CCACCGCAGT AAAGACTAGG TAAACTGGTC ACAGTTAGCA TTTCGGGTCG

2761 TTATAAGTTT TCTATAATAT AATATGTTGC ACTTTTTAGT ATAGTGTGAT TTTATCACCT

2821 TTGTACTGTT TATGTTAGTG TGGTTTAACC ATCTTAGTGT GCTTTATATT ATAGTTTATG

2881 CATAACAGGG AGAACCATTA CAATACCGGA GTTGTTTGTA GTGTGATTAC ATCACGGTTA

2941 ATAGCCGAGG TACGGTAATG TTTGTTGTCC TAAAA

FIGURE 5 bis (a)

```
  1   GAATTCTAGA TTCTGGAAAC ATTGGATAAG   CACTCAAGAT GGTTTTGTG GGCTACCTCT
 61   CGTGTCTGTC AATGATGGAT ATATTGTTGG   TATACATGGA TTGACATCTA ATGATTCAGA
121   GAAGAACTTC TTTGTTCCAT TTACTGACGG   ATTTGAAACG GAATACCTGA ACAATGCTGA
181   TAATCTGTCG TGGGACAAAC ACTGGTTTTG   GGAACCAAGT AAGATCGCAT GGGGCTCTTT
241   GAATTTAGTA GAAGAACAGC CAAAGGAAGA   ATTCAAGATA TCAAAACTCG TGTCTGATCT
301   TTTTGGCAAC ACGGTAGCAG TACAGAGTAG   AAAAGAGAGA TGGGTTTTGG ATGCCATGGA
361   AGGGAATTTG ATAGCTTGTG GGCAAGCTGA   TAGCGCACTA GTGACAAAGC ACGTGGTTAA
421   AGGAAAATGC CCTCATTTTG CACAATACCT   TTCGTTACAT GATGAAGCAA AACAGTTCTT
481   TGAGCCATTG ATGGGGGCAT ACCAGCCAAG   TCGGTTGAAT AAAGACGCTT TCAAGAAAGA
541   TTTCTTTAAA TACAACAAAC CGGTTGTTCT   GAATGAAGTT GATTTTAATG CTTTCGAGAA
601   GGCAGTCGAG GGAGTGGTAA CAATGATGGT   CGACTTTGAA TTTGCAGAAT GCTTATTGT
661   GACTGATCCT GATGAGATTT ATGGATCCCT   GAACATGAAA GCTGCTGTCG GTGCTCAATA
721   CAAAGGAAAG AAACAGGATT ACTTTTCTGG   CATGGATAGT TTTGATAAGG AGCGCCCTGTT
```

FIGURE 5 bis (b)

```
 781  GTACCTTAGC TGTGAGAGGC TTTTTAACGG GGAGAAAGGA ATTTGGAATG GTTCCTTGAA
 841  AGCGGAGTTG AGGCCAATTG AAAAAGTGCA GGCGAACAAA ACTAGAACAT TTACAGCAGC
 901  ACCAATTGAC ACATTGCTTG GAGCTAAAGT TTGCGTCGAT GATTTAAATA ATCAATTCTA
 961  TAGCTTCAAC TTAAAATGTC CATGGACAGT TGGCATGACT AAGTTTTATG GGGGTTGGGA
1021  TAAGCTAAATG AGGAGTTTGC CTGATGGCTG GACATATTGC CACGCAGACG GTTCACAGTT
1081  CGACAGTTCT TTGACTCCCT TATTGTTGAA TGCTGTCCTC AGTATTAGAT GTTGCTTCAT
1141  GGAAGATTGG TGGGTCGGAA AGGAAATGCT CGAAAACCTT TACGCTGAGA TAGTTTACAC
1201  ACCAATCTTA GCACCTGATG GCACAATTTT CAAAAAGTTT AGAGGGAATA ACAGTGACA
1261  GCCATCTACA GTTGTTGACA ATACGCTTAT GGTTGTCATT GCTATGTATT ACTCGTGCTG
1321  TAAGCAAGGT TGGTCAGAGA AGGATATTGA AGAAAGGTTG GTGTTCTTTG CTAATGGTGA
1381  TGATATTATC TTGGCAGTTA GGGACGAGGA TGTGTGGCTG TATGACACTC TAAGTGCTTC
1441  ATTCGCTGAG CTTGGCCTGA ATTATAAACTT CGATGAGCGG ACAAAGAAAA GAGAGGAGCT
1501  ATGGTTTATG TCACACCAAG CCATGCTAGT TGATGGGATT TATATTCCAA AGCTTGAACC
1561  TGAAAGGATA GTTTCCATCT TAGAGTGGGA TAGGAGCAAG GAACTTATGC ATCGCACTGA
```

FIGURE 5 bis (c)

```
1621  GGCAATTTGT  GCAGCCATGA  TTGAGGCATG  GGGACATACT  GAGCTTTTGC  AGGAAATTCG
1681  CAAATTTTAT  TTGTGGCTTC  TAAGCAAAGA  TGAATTTAAG  GAACTCGCTG  CATCCGGAAA
1741  AGCTCCATAC  ATAGCTGAGA  CAGCATTGAG  AAAGCTCTAC  ACAGATGTTA  ACACACAACC
1801  AAGTGAATTG  CAAAAATACC  TTGAAGTTTT  GGATTTTAAC  CACATTGATG  GTTGTTGTGA
1861  ATCAGTGTCT  CTGCAATCAG  GAAAAGAAGC  AGTAGAAAAT  TTGGATGCAG  GGAAGGACTC
1921  GAAGAAGGAC  ACCAGTGGCA  AAGGGGATAA  GCCACAAAAC  TCGCAAACTG  ATCAAGGTAG
1981  CAAAGAACAG  ACAAAAACTG  GCACAGTCAG  CAAAGATGTG  AATGTTGGAT  CGAAAGGAAA
2041  AGAAGTCCCA  CGACTACAAA  AGATAACAAA  GAAAATGAAC  CTTCCAACAG  TTGGTGGGAA
2101  AATCATTCTT  AGCTTAGACC  ATTTGCTCGA  GTACAAACCT  AATCAAGTTG  ATCTGTTTAA
2161  CACTCGAGCA  ACAAAAACAC  AGTTTGAGTC  ATGGTACAGC  GCAGTTAAAG  TTGAATATGA
2221  TCTTAACGAT  GAGCAAATGG  GTGTGATTAT  GAATGGTTTT  ATGGTTTGGT  GTATCGATAA
2281  CGGTACATCT  CCAGATGTCA  ATGGAGTTTG  GGTGATGATG  GATGGGGAAG  AGCAAGTTGA
2341  GTATCCATTA  AAGCCAATTG  TTGAAAATGC  AAAACCAACT  TTAAGACAAA  TCATGCACCA
2401  TTTCCTCAGAC  GCAGCAGAAG  CATATATTGA  AATGAGAAAC  TCTGAAAGTC  CGTATATGCC
```

FIGURE 5 bis (d)

```
2461 TAGATACGGA TTACTAAGAA ATTTGAGAGA CAGGGAATTA GCACGCTATG CTTTTGACTT

2521 TTATGAGGTT ACTTCCAAAG CACCTAATAG GGCAAGAGAA GCAATAGCAC AAATGAAGGC

2581 CGCAGCTCTC GCGGGAATTA ACAGCAGGTT ATTTGGGCTT GATGGTAATA TCTCGACCAA

2641 TTCCGAAAAT ACTGAGAGGC ACACTGCAAG GGACGTGAAT CAGAATATGC ATACTTTGTT

2701 AGGTATGGGT CCACCGCAGT AAAGACTAGG TAAACTGGTC ACAGTTAGCA TTTCGGGTCG

2761 TTACAAGTTT TCTATAATAT AATATGTTGC ACTTTTTAGT ATAGTGTGAT TTTATCACCT

2821 TTGTTACTGTT TATGTAGTGT GGTTTAACCA CCTTAGTGTG CTTTATATTA TAGTTTATGC

2881 ATAACAGGGA GAACCATTGC AATACCGGAG TTGTTTGTAG TGTGATTACA TCACGTTTAA

2941 TAGCCGAGGT ACGGTAATGT TTGTTGTCCT AAAA
```

FIGURE 6(a)

```
  1 TCTAGATTCT GGAAACATTG GATAACNACC AACGAAGGGG ATTGCGGATT GCCCATGGTT  60
 61 TCAACAACGG ATGGCAAGAT AATTGGAGTT CATGGTTTGG CTTCCACAGT CTCATCTAAG 120
121 AATTATTTTG TCCCATTCAC TGATGATTTT ATAGCCACGC ATTTGAGCAA ACTTGATGAC 180
181 CTCACATGGA CTCAGCATTG GCTATGGCAA CCTAGCAAAA TTGCGTGGGG AACGCTCAAC 240
241 TTAGTTGATG AACAACCAGG GCCCGAATTT CGTATTCAA  ATCTAGTCAA GGATTATTC  300
301 ACTTCTGGTG TTGAAACACA GAGCAAGCGA GAAAGATGGG TCTACGAAAG CTGTGAAGGG 360
361 AACCTTCGGG CTGTTGGAAC TGCACAATCA GCGTTAGTCA CCAAACATGT TGTGAAAGGC 420
421 AAGTGTCCTT TCTTCGAAGA ATATTTACAA AAACACGCAG AAGCGGAGCGC CTATTTCAGA 480
481 CCCCTAATGG GAGAGTACCA GCCGAGCAAG TTGAACAAAG AAGCCTTTAA AAAGGATTTC 540
541 TTTAAATACA ATAAACCCGT CACTGTTAAT CAACTGGATC ATGATAAATT TTTGGAAGCA 600
601 GTGGATGGGG TTATACGTAT GATGTGTGAC TTTGAGTTCA ACGAATGTCG ATTCATTACA 660
661 GATCCCGAGG AGATTTACAA CTCTTTGAAC ATGAAAGCAG CAATTGGAGC CCAGTATAGA 720
721 GGAAAGAAGA AAGAGTNTTT TGAGGGGCTA GATGATTTTG ATCGAGAGCG ACTTTTATTC 780
```

FIGURE 6 (b)

```
 781  CAAAGTTGTG AAAGGTTGTT CAATGGCTAC AAAGGTCTGT GGAATGGATC TTTAAAGGCC  840
 841  GAGCTCAGGC CGCTTGAGAA AGTCAGGGCT AACAAAACAC GAACCTTTAC AGCAGCGCCA  900
 901  ATTGATACAT TGCTTGGAGC TAAAGTTTGT GTGGATGATT TCAACAATGA GTTCTACAGG  960
 961  AAAAACCTCA AGTGTCCATG GACGGTCGGC ATGACAAAAT TTTATGGTAG TTGGGATAAA 1020
1021  TTGATGAGAT CATTACCTGA TGGTTGGTTG TATTGTCATG CTGATGGATC ACAGTTCGAT 1080
1081  AGTTCGTTAA CCCCAGCCTT ACTGAACGCA GTGCTCATAA TCAGTCATT TTATATGGAG 1140
1141  GATTGGTGGG TTGGCCAAGA GATGCTTGAA AATCTTTATG CCGAGATTGT GTACACTCCA 1200
1201  ATTCTTGCTC CTGATGGAAC AATTTTCAAG AAATTTAGAG GTAACAACAG TGGGCAACCC 1260
1261  TCAACAGTGG TGGATAAACAC ACTAATGGTT GTGATCTCTA TTTACTATGC GTGCATGAAA 1320
1321  TTTGGTTGGA ACTGCGAGGA GATTGAGAAT AAACTTGTCT TCTTTGCAAA TGGGGATGAT 1380
1381  CTGATACTTG CAGTCAAAGA TGAGGATAGC GGCTTACTTG ATAACATGTC ATCCTCTTTT 1440
1441  TGCGAACTTG GACTGAATTA TGATTTTTCA GAACGTACGC ATAAAAGAGA AGATCTTTGG 1500
1501  TTCATGTCCC ACCAAGCAAT GCTAGTTGAT GGAATGTACA TTCCAAAACT CGAGAAAGAG 1560
```

FIGURE 6 (c)

```
1561 AGAATTGTTT CAATTCTAGA GTGGGATAGA AGCAAAGAAA TTATGCACCG AACAGAGGCT 1620
1621 ATTTGCGCTG CGATGATTGA GGCATGGGGG CACACCGAGC TCTTGCAAGA AATCAGAAAG 1680
1681 TTTTACCTAT GGTTCGTTGA AAAAGAAGAA GTGCCGAGAAT TGGCAGCCAT CGGGAAAGCT 1740
1741 CCATACATAG CTGAGACAGC ACTTCGTAAG TTATACACTG ACAAGGGAGC AGATACAAGT 1800
1801 GAACTGGCAC GCTACCTACA AGCCCTCCAT CAAGACATCT TCTTTGAACA AGGAGACACT 1860
1861 GTGATGCTCC AATCAGGCAC TCAGCCAACT GTGGCAGATG CTGGAGCTAC AAAGAAAGAC 1920
1921 AAAGAAGATG ACAAAGGGAA AAACAAGGAC GTTACAGGCT CTGGCTCAGG TGAGAAAACA 1980
1981 GTAGCAGCTG TCACGAAGGA CAAGGATGTG AATGCTGGTT CTCATGGGAA AATTGTGCCG 2040
2041 CGTCTTTCGA AGATCACAAA GAAAAATGTCA TTGCCACGCG TGAAAGGAAA TGTGATACTC 2100
2101 GATATTGATC ATTTGCTGGA ATATAAACCG GATCAAAATTG AGTTATATAA CACACGAGCG 2160
2161 TCTCATCAGC AGTTCGCCTC TTGGTTCAAC CAGGTTAAGA CGGAATATGA TTTGAACGAG 2220
2221 CAACAGATGG GAGTTGTAAT GAATGGTTTC ATGGTTTGGT GTATTGAAAA TGGCACTTCA 2280
2281 CCCGACATTA ATGGAGTGTG GGTTATGATG GACGGAAATG AGCAAGTTGA GTATCCCTTG 2340
```

FIGURE 6 (d)

```
2341 AAACCAATAG TTGAAAATGC AAAGCCAACG CTGCGGCAAA TAATGCATCA TTTTTCAGAT 2400
2401 GCAGCGGAGG CATATATAGA GATGAGAAAT GCAGAGGCAC CATACATGCC GAGGTATGGT 2460
2461 TTGCTTCGAA ACCTACGGGA TAGGAGTTTA GCACGATATG CTTTCGATTT CTATGAAGTC 2520
2521 AATTCTAAAA CTCCTGAAAG AGCCCGCGAA GCTGTTGCGC AGATGAAAGC AGCAGCTCTT 2580
2581 AGCCAATGTT CTTCAAGGTT GTTTGSCCTT GATGGAAATG TTGCCACCAC TAGTGAAGAC 2640
2641 ACTGAACGGC ACACTGCACG TGATGTTAAT AGAAACATGC ACACCTTACT AGGTGTGAAT 2700
2701 ACAATGCAGT AAAGGGTAGG CCGCCTACCT AGGTTATTGT TTCGCTGCCG ACGTAATTCT 2760
2761 AATATTTACC GCTTTATTTG ATATCTTTAG ATTTCCAGAG TGGGCCTCCC ACCTTTAAAG 2820
2821 CGTAAAGTTT ATGTTAGTTG TCCAGGAGTG CCGTAGTCCT GTCGGAAGCT TTAGTGTGAG 2880
2881 CCTCTCACGA ATAAGCTCGA GATTAGACTC CGTTTGAAGC CTAAA
```

FIGURE 7A

gagctccacc gcggtggcgg ccgctctaga actagtggat ccatgtcatT
        BstXI   NotI   BamHI

TTCTTttcgc gtcaccgctc atcagGTTAT CTTTTGTCAG CTATGTttcg
        Ribozyme 2067

cgtcaccgct catcagTGGA GCTTTTCCag atctATAAAA CTTAGTttcg
 Ribozyme 1747

cgtcaccgct catcagATGC CAACTGCAGC AGCTTTttcg cgtcaccgct
 Ribozyme 996   Ribozyme 696

catcagATGT TCAGGGAcct gcatctgatc aagctt
                        HindIII

FIGURE 7B

gagctccacc gcggtggcgg ccgctctaga actagtggat ccATGAATGG

        BstXI   NotI   BamHI

TTTttcgcgt caccgctcat cagATGGTTT GGTGCACCAT TTttcgcgtc

        Ribozyme 2259

accgctcatc agTCAGACGC AGatCTTTTG ACTTttcgcg tcaccgctca

Ribozyme 2403   Ribozyme 2520

tcagTATGAG GTTACCACCG CAGTttcgcg tcaccgctca tcagAAGACT

                                Ribozyme 2720

AGGTAAcctg catctgatca agcttgaatt

        BspMI   HindIII

FIGURE 7C

```
     aagcttgatatcg   2678 AAT CAGAATATGC ATACTTTGTT
        EcoRV  (EcoRI)-HInfI

2701 AGGTATGGGT CCACCGCAGT AAAGACTAGG TAAACTGGTC ACAGTTAGCA
2751 TTTCGGGTCG TTACAAGTTT TCTATAATAT AATATGTTGC ACTTTTTAGT
2801 ATAGTGTGAT TTTATCACCT TTGTACTGTT TATGTAGTGT GGTTTAACCA
2851 CCTTAGTGTG CTTTATATTA TAGTTTATGC ATAACAGGGA GAACCATTGC
2901 AATACCGGAG TTGTTTGTAG TGTGATTACA TCACGTTTAA TAGCCGAGGT
2951 ACGGTAATGT TTGTTGTCCT AAAAAAAAAA ggatccacta gttctagagc
                                        BamHI

ggccgccaccgc
   NotI
```

Figure 8A

gcggccgct ctagaactag tggatccatG CAATGTTTtt cgcgtcaccg
    NotI   BamHI   Ribozyme 2590

ctcatcagTT CAAGGTGCCG AGGTttcgcg tcaccgctca tcagTGGTTT
                              Ribozyme 2455

GCTTCATGGT ttcgcgtcac cgctcatcag TGGTGCATAC GAGCGTttcg
            Ribozyme 2255

cgtcaccgct catcagTCAT CAGCTCCACA GTttcgcgtc accgctcatc
Ribozyme 2161   Ribozyme 110

agTCATCTAA ctgcatctga tcaagctt
                      HindIII

FIGURE 8B

aagcttgatc agatgcagGC TGCTTTttcg cgtcaccgct catcagATCT

HindIII   Ribozyme 2564

GCGCCGGCAT GTttcgcgtc accgctcatc agTGGTGCCT CCCTTTGTtt

                Ribozyme 2442

cgcgtcaccg ctcatcagAT CTTCTTAGCT ATGTttcgcg tcaccgctca

 Ribozyme 1929   Ribozyme 1743

tcagTGGAGC TTCCCACTGT ttcgcgtcac cgctcatcag GTTACCTCat

                Ribozyme 1245

ggatccacta gttctagagc ggccgc

 BamHI   **NotI**

FIGURE 8C

ggatccatGC TTTGCATTtt cgcgtcaccg ctcatcagTC AACTATcGAT

 BamHI   Ribozyme 2354

GCCGAGGTtt cgtcctcacg gactcatcag TGGTTTGCTT ttcgtcctca

           Ribozyme 2455   Ribozyme

cggactcatc agGAAACCTA CGctgcAGCA ATGTTTttcg tcctcacgga

 2466   Ribozyme 2590

ctcatcagTT CAAGGTTGTT ttcgcgtcac cgctcatcag GGCCTTGATG

                        Ribozyme 2603

acctgcatct gatcaagctt

          HindIII

FIGURE 8D

ggatccGCAC AATTTTttcg cgtcaccgct catcagCCAT GAGAACCAGC

BamHI   Ribozyme 2027

ATTttcgcgt caccgctcat cagACATCCT TGTgatc**aag ctt**

      Ribozyme 2009   **HindIII**

FIGURE 9 (a)

TTATTTGCCGCAGCGTTGGCTTTGCATTttcgtcctcacggactcatcagTCAACTATTG
*base 2383   Ribozyme 2354(-)

GTTTCAAGGGATACTCAACTTGCTCATTTCCGTCCATCATAACCCACACTCCATTAATGT

CGGGTGAAGTGCCATTTTCAATACACCAAACCATTGAAACCATTCATTACAACTCCCATCT

GTTGCTCGTTCAAATCATATTCCGTCTTAACCTGGTTGAACCAAGAGGCGAACTGCTGAT

GAGACGCTCGTGTGTTATATAACTCAATTTGATCCGGTTTATATTCCAGCAAATGATCAA

TATCGAGTATCACATTTCCTTTCACGCGTGGCAATGACATTTCTTTGTGATCTTCGAAA

GACGCGGCACAATTTTtcgtcctcacggactcatcagCCATGAGAACCAGCATTtcgt
                        Ribozyme 2027(-)

cctcacggactcatcagACATCCTTGCCCTTCGTGACAGCTGCTACTGTTTCTCACCTG
Ribozyme 2009(-)

FIGURE 9 (b)

AGCCAGTCATTTTCAGATGCAGCGGAGGCATATATAGAGATGAGAAATGCAGAGGCACC
b1961**base 2388

ATACATGCCGAGGTttcgtcctcacggactcatcagTGGTTTGCTTttcgtcctcacgga
         Ribozyme 2455 (+)   Ribozyme 2466 (+)
ctcatcagGAAACCTACGGGATAGGAGTTTAGCACGATATGCTTTCGATTTCTATGAAGT

CAATTCTAAAACTCCTGAAAGAGCCCGCGAAGCTGTTGCCGAGATGAAAGCAGCAGCTCT

TAGCAAATGTTTttcgtcctcacggactcatcagTTCAAGGTTGTTttcgtcctcacggac
       Ribozyme 2590 (+)   Ribozyme 2603 (+)
tcatcagGGCCTTGATGGAAATGTTGCCACCACTAGTGAAGACACTGAACGGCACACTGC

ACGTGATGTTAATAGAAACATGCACACCTTACTAGGTGTGAATACAATGCAGTAAAGGGT

AGGCCGCCTACCTAGGTTATTGTTTCGCTGCCGACGTAATTCTAATATTTACCGCTTTAT

TTGATATCTTTAGATTTCCAGAGTGGGCCTCCCACCTTTAAAGCGTAAAGTTTATGTTAG

TTGTCCAGGAGTGCCGTAGTCCTGTCGGAAGCTT
              base 2871* (HindIII)

Figure 10

Figure 11

MW	-	+	+	+	+	+	+	+	: Ribozyme
	7.30	0.45	1.30	3.30	3.30	7.30	7.30	24	: Time h.mn
markers	37	50	50	50	37	37	22	22	: Temp. ºC

Bands:

2 - 3'
5' - 4

3 - 3'

5' - 3

2 - 4
4 - 3'

5' - 2

(3 - 4 ran out of the gel)

5' ———————————————— 3' Target

| 2259 | 2403 | 2520 | 2720 | | Sites |
| 1 | 2 | 3 | 4 | | |
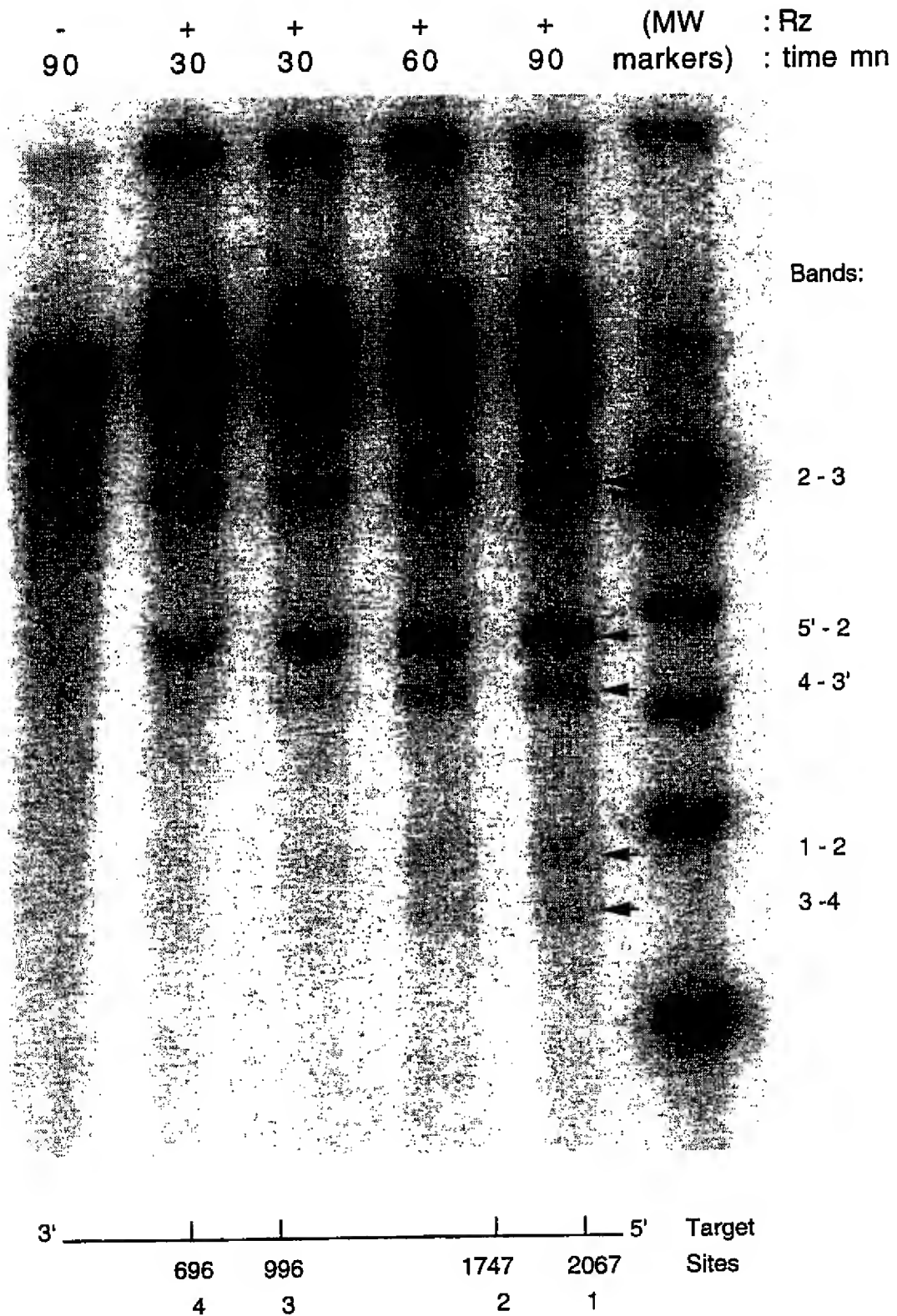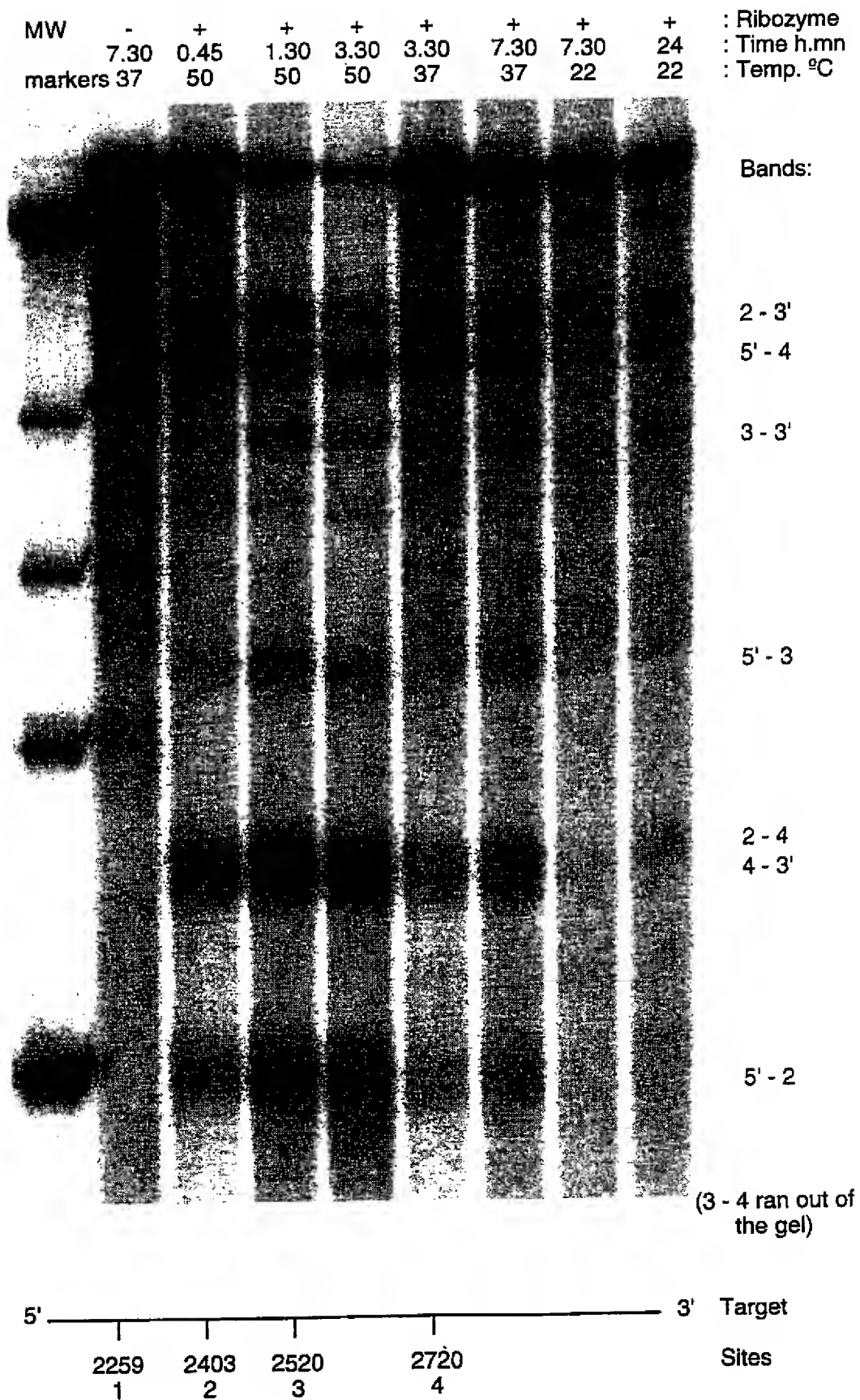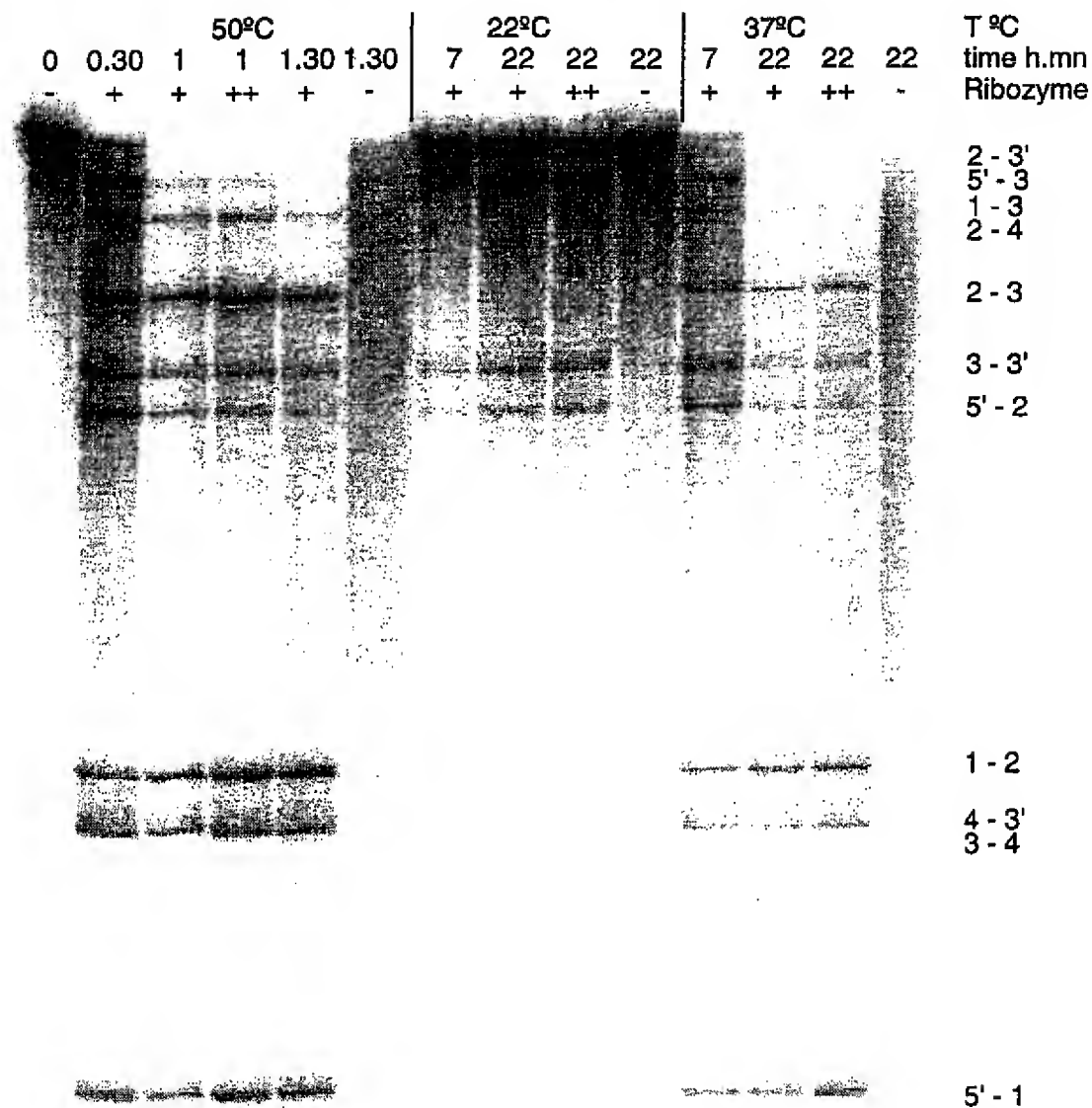
Figure 12

Ribozyme   - : no ribozyme
                  + : 4 fold excess
                  ++: 8 fold excess
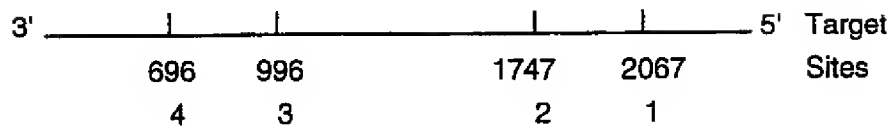
Figure 13
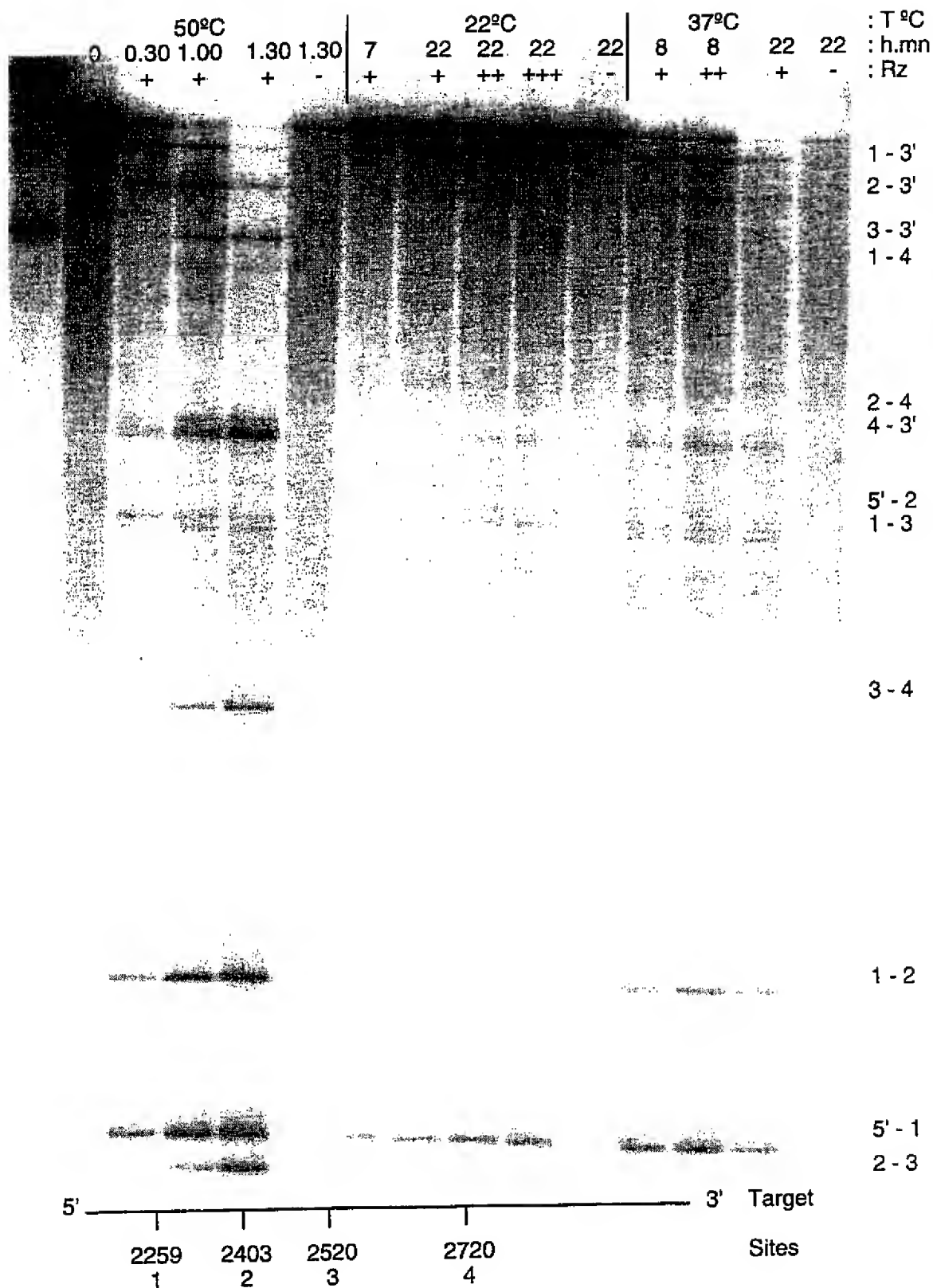
Figure 14
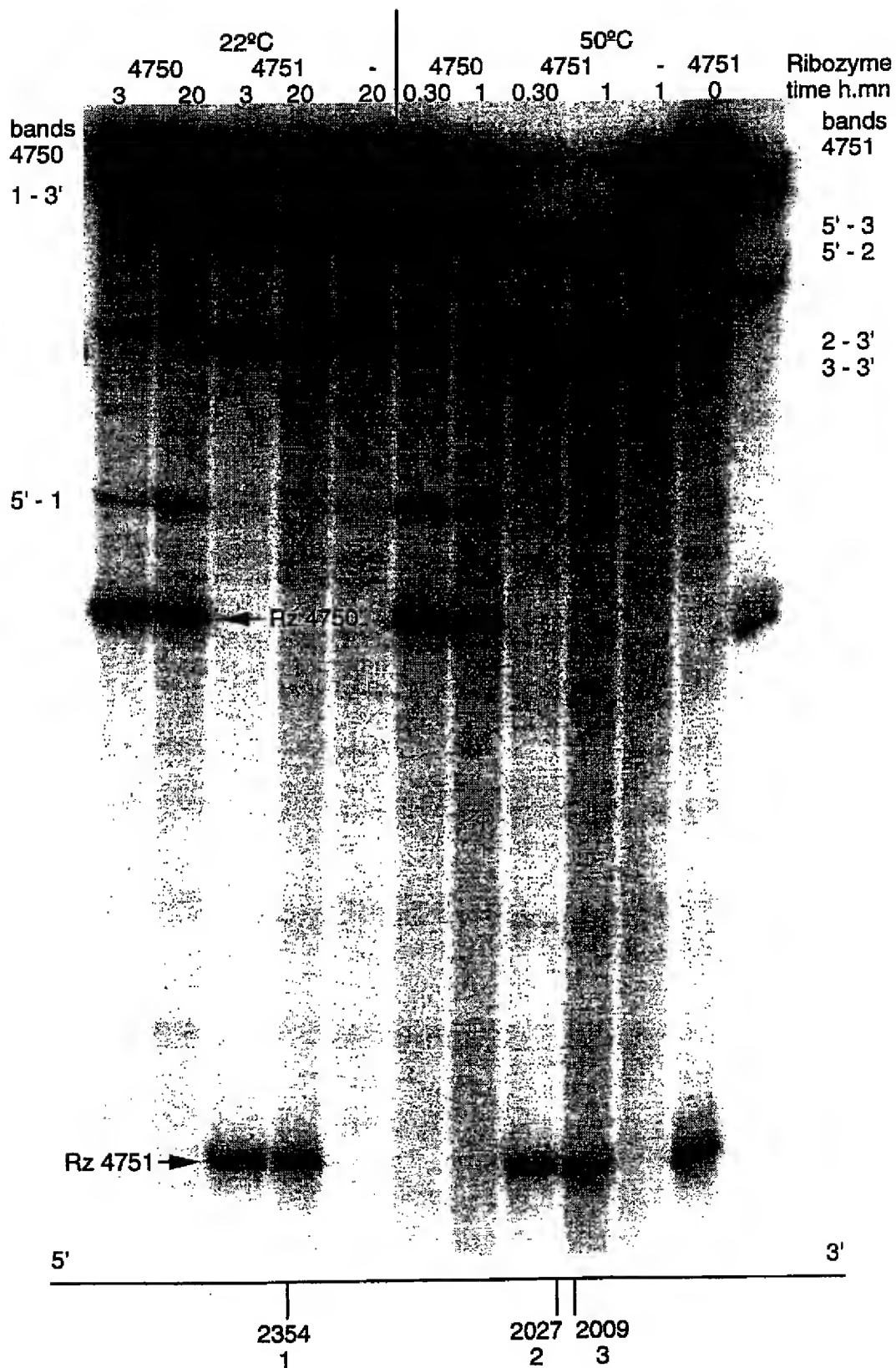
(A) Target plasmids
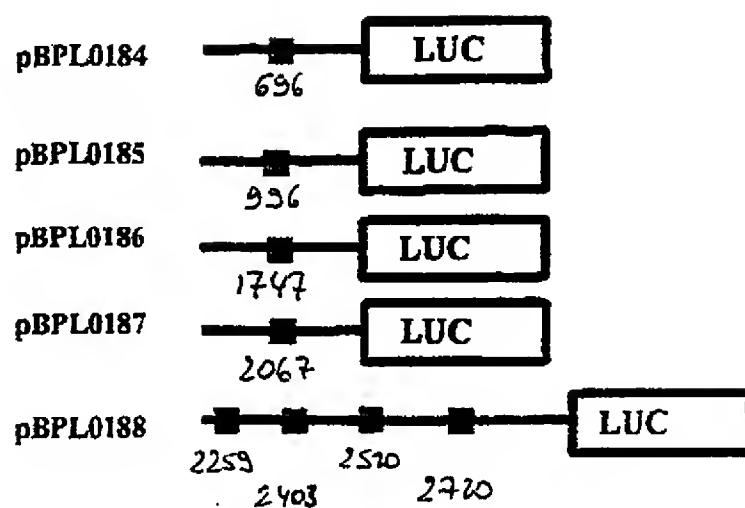
(B) Ribozymes
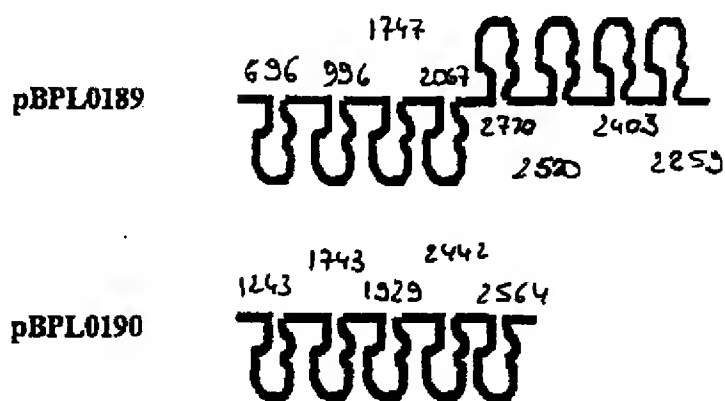
Figure 15
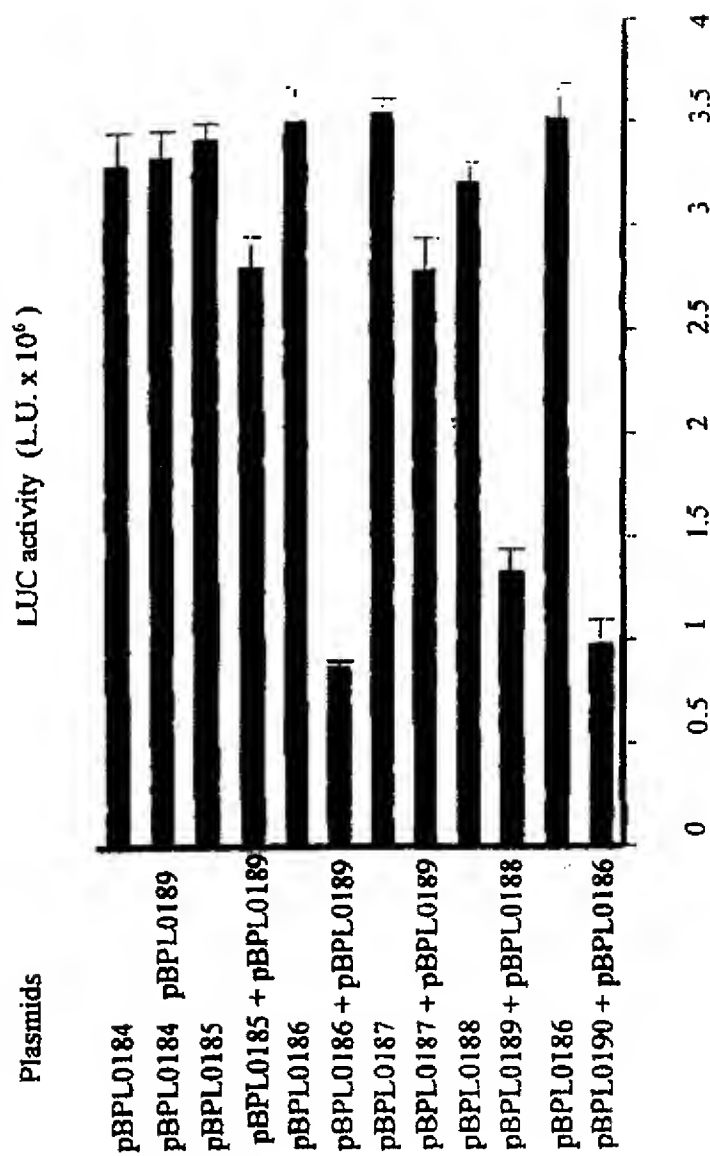
Figure 16
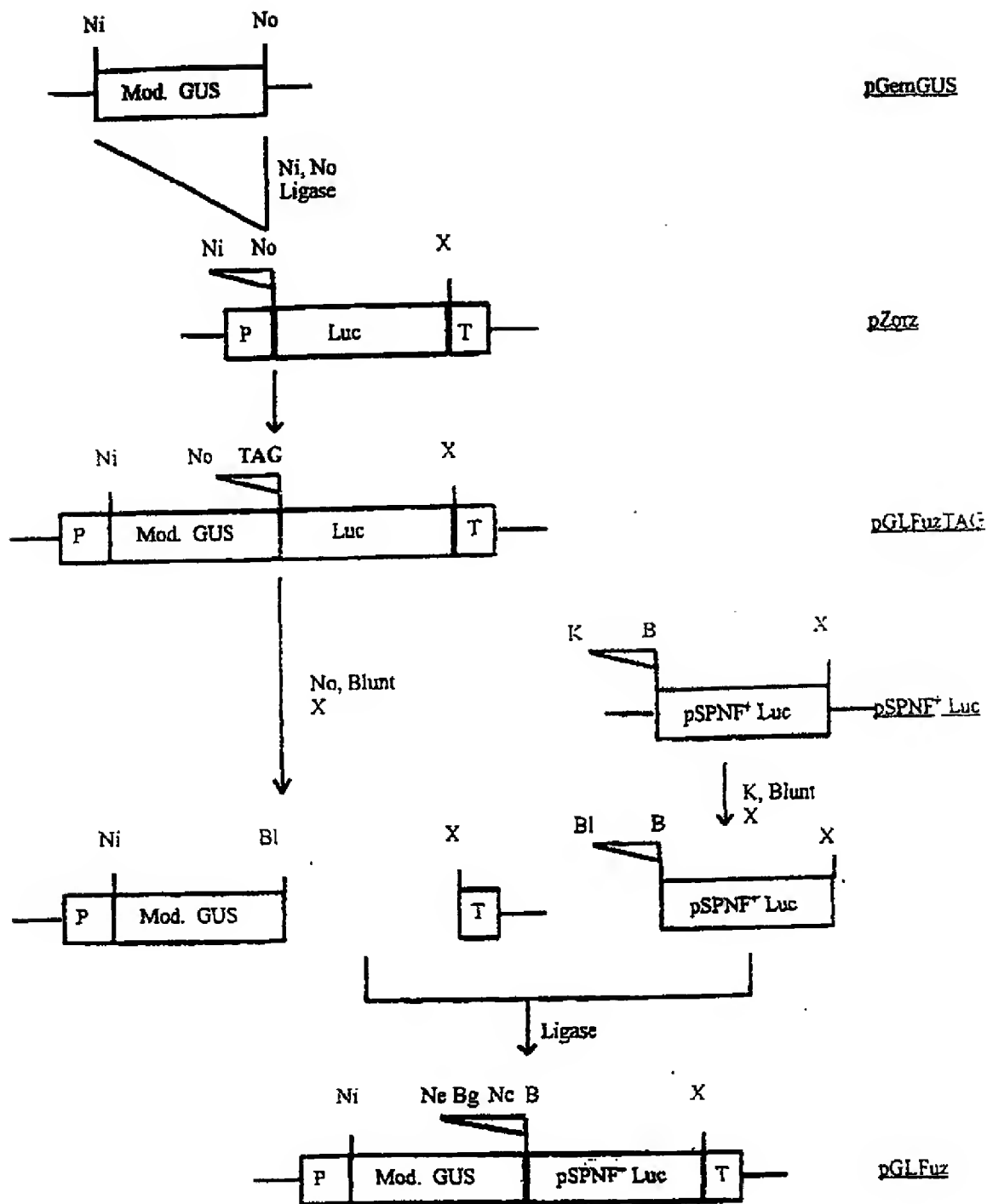
Figure 17

5'

```
(+) CAAATTTTAT  TTGTGGCTTC  TAAGCAAAGA  TGAATTTAAG  CATCCGGAAA
(-) GTTTAAAATA  AACACCGAAG  ATTCGTTTCT  ACTTAAATTC  CTTGAGCGAC  GTAGGCCTTT
    1681

(+) AGCTCCATAC  ATAGCTGAGA  CAGCATTGAG  AAAGCTCTAC  ACACACAACC
(-) TCGAGGTATG  TATCGACTCT  GTCGTAACTC  TTTCGAGATG  TGTCTACAAT  TGTGTGTTGG
    1741        ↑

(+) AAGTGAATTG  CAAAAATACC  TTGAAGTTTT  GGATTTTAAC  CACATTGATG  GTTGCTGTGA
(-) TTCACTTAAC  GTTTTTATGG  AACTTCAAAA  CCTAAAATTG  GTGTAACTAC  CAACGACACT
    1801

(+) AT(C/A)AGTGTCT  CTGCAATCAG  GAAAAGAAGC  AGTAGAAAAT  TTGGATGCAG  GGAAGGACTC
(-) TA(G/T)TCACAGA  GACGTTAGTC  CTTTTCTTCG  TCATCTTTTA  AACCTACGTC  CCTTCCTGAG
    1861
```

FIGURE 18

POLYRIBOZYME pBL3817

TATCAAAGCTCGTGTCTGATCTTTTTGGCAACACGGTAGCAGTAGAAAAGAGAGATGGGTTTTGGACGCCATG.
GAAGGGAATTTGATAGCTTGTGTGGCAAGCTGATAGCGCACTAGTGACAAAGCACGTGGTTAAAGGAAAATGCCCTCATTT
TGCACAATATCTTTCGTTACATGATGAAGCAAAACAGTTCTCCTGAGCCATTGATGGGGCATACCAGCCAAGTCGGTTGA
ATAAAGACGGCTTCAAGAAAGATTTCTTTAAGTACAACAAACCGGTTGTCTGAATGAAGTTGATTTAAATGCTCTCGAG
AAGGCAGTCGAGGGAGTGATAACAATGATGGTCGACTTTGAATTTGCAGAATGCTTATTTGTGACTGATCCTGATGAGAT
TTATGGATCCCTGAACATCTGATGAGTCCCGTGAGGACGAAAAAGCTGCTGTCGGGTGCTCAATACAAAGGAAAGAAACAG

GATTACTTTCTTGGCATGGATAGTTTTGATAAGGAGCGCCTGTTGTGTACCTTAGCTGTGAGAGGCTTTTAATGGAGGAA
AGGAATTGGAACGGGTTCCTTCAAACGGGAGTTGAGGCCCAATTGAAAAGTGCAGCCAACAAAACTAGAACATTTACAG
CAGCACCGATTGATACATTGCTTGGAGCTAAAGTTTGCGTCGATGATTTAATAATCAATTCTACAGCTTCAACTTAAAA
TGTTCCATGACAGTTGGATCTGATGAGTCCTGAGGACGAAACTAAGTTTTATGGGGGTTGGGATAAGCTAATGAGGAGTT
TTGGCCTGATGGGCTGGACATATTGCCACGCAGACGGTTCACAGTTCTTGACTCCCTTATTGTTGTGAATGCTGT
CCTCAGTAGTAGATGTTGCTTCATGGAGAAGATTGGTGGGTTGGAAAGGAAAATGCTTGAAAACCTTTACCGCTGAGATAGTTT
ACACACCAATCTTAGCCGATGGCCACAATTTCAAAAAGTTTAGAGGGAATAACAGTGGAATAACAGTGGACCAGCCATCTACAGTTGTT
GACAATACGGCTTATGGTTGTCATTGCTATGGTGATGATTATCTTGGCAGTTAGGGATGAGGATGTGTGGCTGTATGACACTCTAAGTG
CTTCATTCGGCTGAGCTTGGGCCTGAATTATAACTTCCAAAGCTTGAACCTGAAAGGATAGTTTCCATCTTAGAGTGGGATTGGAG
CAAGCCATGCTAGTTGATGGGAATTTACATTCCAAATTTGTGCAGCCAATTTGTGAGGCATGGGGACATACTGAGCTTTTGCAGGAAA
TTCGCAAATTTTATTTGTGGCTTCTAAGCAAAGATGAATTTAAGGAACTCCGTCGCATCCGGAAAAAGCTCCACTGATGAG
TCCCGTGAGGACGAAACATAGCTGAGACAGCATTGAGAAGCTCTACACACAGATGTTAACACACAAGTGAATTGCAAA
AATACCTTGAAGTTTTGGATTTTAACCACATTGATGGTTGCTGTGAATCAGTGTCTCTGCAATCAGGAAAGAAGCAGTA
GAAAATTTGGATGCAGGGAAGGACTCGAAGAAGGACACCGAGTGCAAAGGGGATAAGCCACAAAATTCGCAAACTGGTCA
AGGTAGCAAAGAACAGACARAAACTGGTACAGTGCAAAGATGTGAATGTTTGGATCGAAAGGAAAGARGTCTCACGAC
TACAAAGATAACCTGATGAGTCCGTGAGGACGAAATGAAACCTTCCACACGTTGGTGGGAAAATCATTCTTAGC
TTAGCCATTTACTCGAGTACAAACCTAATCAAGTTGATCTTTAACACTCGAGCACTCGAGCCCAGCTTCCACCATGGCGTGCAGGTCG
ACGGA
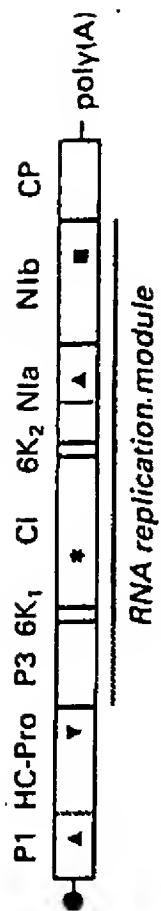
Figure 19

FIGURE 20

FIGURE 21

Epidemiology in transgenic plant field trial (Season N° 2)
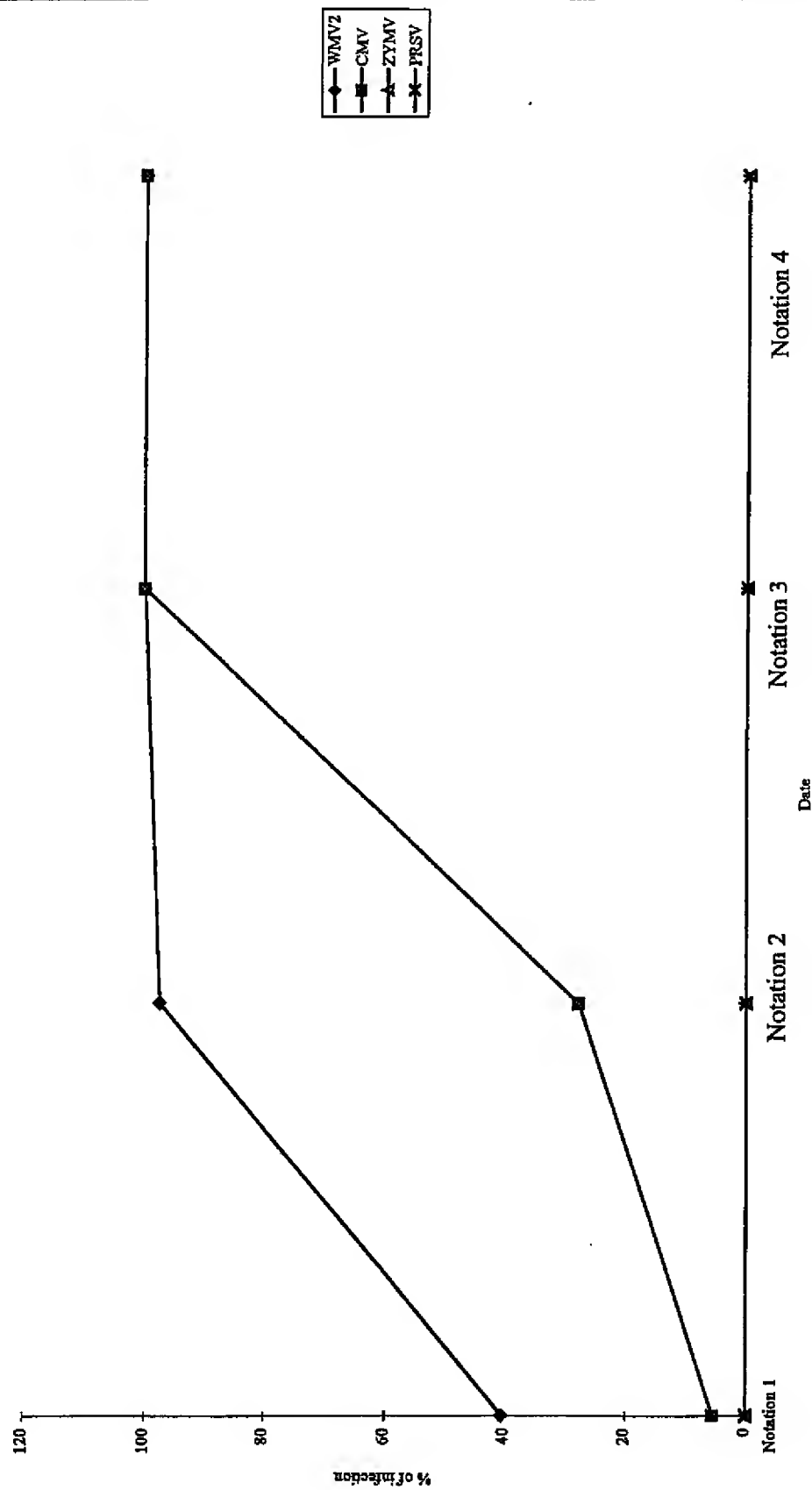
SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT:
        (A) NAME: GENE SHEARS PTY LTD
        (B) STREET: Suite 1, Building 5 - 105 Dehli Road
        (C) CITY: NORTH RYDE
        (E) COUNTRY: AUSTRALIA
        (F) POSTAL CODE (ZIP): NSW 2113

    (ii) TITLE OF INVENTION:
        RIBOZYMES CAPABLE OF CONFERRING RESISTANCE TO POTYVIRUS INFECTION,
        AND PLANTS EXPRESSING SAID RIBOZYMES.

    (iii) NUMBER OF SEQUENCES: 109

    (iv) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

    (v) CURRENT APPLICATION DATA:
        APPLICATION NUMBER: PCT/EP

(2) INFORMATION FOR SEQ ID NO: 1:

 (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 21 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

 (ii) MOLECULE TYPE: other nucleic acid
  (A) DESCRIPTION: /desc = "SYNTHETIC RNA"

 (ix) FEATURE:
  (A) NAME/KEY: misc_feature
  (B) LOCATION:1..21
  (D) OTHER INFORMATION:/note= "Mutant ribozyme I"

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CUGAUGAGUC CUGAGGACGA A   21

(2) INFORMATION FOR SEQ ID NO: 2:

 (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 23 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

 (ii) MOLECULE TYPE: other nucleic acid
  (A) DESCRIPTION: /desc = "Synthetic RNA"

 (ix) FEATURE:
  (A) NAME/KEY: misc_feature
  (B) LOCATION:1..23
  (D) OTHER INFORMATION:/note= "Mutant ribozyme II

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CUGAUGAGUC CCGUGAGGAC GAA   23

(2) INFORMATION FOR SEQ ID NO: 3:

 (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 21 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

 (ii) MOLECULE TYPE: other nucleic acid
  (A) DESCRIPTION: /desc = "SYNTHETIC RNA"

 (ix) FEATURE:
  (A) NAME/KEY: misc_feature
  (B) LOCATION:1..21

          (D) OTHER INFORMATION:/note= "Mutant ribozyme III"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CUGAUGAGUC GUGAGGACGA A   21

(2) INFORMATION FOR SEQ ID NO: 4:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 20 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: other nucleic acid
          (A) DESCRIPTION: /desc = "SYNTHETIC RNA"

     (ix) FEATURE:
          (A) NAME/KEY: misc_feature
          (B) LOCATION:1..20
          (D) OTHER INFORMATION:/note= "Mutant ribozyme IV"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CUGAUGAGCG GUGACGCGAA   20

(2) INFORMATION FOR SEQ ID NO: 5:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 22 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: other nucleic acid
          (A) DESCRIPTION: /desc = "synthetic RNA"

     (ix) FEATURE:
          (A) NAME/KEY: misc_feature
          (B) LOCATION:1..22
          (D) OTHER INFORMATION:/note= "Wild type sTobRV"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CUGAUGAGUC CGUGAGGACG AA   22

(2) INFORMATION FOR SEQ ID NO: 6:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 21 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: other nucleic acid
          (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..21
        (D) OTHER INFORMATION:/note= "poly-T-anchor"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GGGAGCTCGC TAGCGGATCC T   21

(2) INFORMATION FOR SEQ ID NO: 7:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 22 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..22
            (D) OTHER INFORMATION:/note= "PRIMER 1"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

AATTGGGAGC TCGCTAGCGG AT   22

(2) INFORMATION FOR SEQ ID NO: 8:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 33 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

        (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..33
            (D) OTHER INFORMATION:/note= "PRIMER 2"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AATTGCTAGC AGTKGTKGAC AATACACTWA TGG   33

(2) INFORMATION FOR SEQ ID NO: 9:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 29 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..29
            (D) OTHER INFORMATION:/note= "primer 3"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AATTCTAGAA THAKATCATC ACCATTKGC   29

(2) INFORMATION FOR SEQ ID NO: 10:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 32 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..32
            (D) OTHER INFORMATION:/note= "Primer 4"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AATTCTAGAT TCTGGAWACA TTGGATAWSN AC   32

(2) INFORMATION FOR SEQ ID NO: 11:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..30
            (D) OTHER INFORMATION:/note= "RBZ696-3'"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GTGAGGACGA AAAAGCTGCT GTCGGTGCTC   30

(2) INFORMATION FOR SEQ ID NO: 12:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 32 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..32
            (D) OTHER INFORMATION:/note= "RBZ696-5'"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGACTCATCA GATGTTCAGG GATCCATAAA TC   32

(2) INFORMATION FOR SEQ ID NO: 13:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 32 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..32
            (D) OTHER INFORMATION:/note= "RBZ996-3'"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GTGAGGACGA AACTAAGTTT TATGGGGGTT GG   32

(2) INFORMATION FOR SEQ ID NO: 14:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (ix) FEATURE:

```
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..30
        (D) OTHER INFORMATION:/note= "RBZ996-5'"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
```

GGACTCATCA GATGCCAACT GTCCATGGAC   30


```
(2) INFORMATION FOR SEQ ID NO: 15:

   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION:   /desc = "SYNTHETIC DNA"

   (ix) FEATURE:                                   .
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..31
        (D) OTHER INFORMATION:/note= "RBZ1747-3'"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
```

GTGAGGACGA AACATAGCTG AGACAGCATT G   31


```
(2) INFORMATION FOR SEQ ID NO: 16:

   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION:   /desc = "SYNTHETIC DNA"

   (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..29
        (D) OTHER INFORMATION:/note= "RBZ1747-5'"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:
```

GGACTCATCA GTGGAGCTTT TCCGGATGC   29


```
(2) INFORMATION FOR SEQ ID NO: 17:

   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 33 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear
```

(ii) MOLECULE TYPE: other nucleic acid
    (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:
    (A) NAME/KEY: misc_feature
    (B) LOCATION:1..33
    (D) OTHER INFORMATION:/note= "RBZ2067-3'"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GTGAGGACGA AAAGAAAATG AACCTTCCAA CAG   33

(2) INFORMATION FOR SEQ ID NO: 18:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 32 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..32
        (D) OTHER INFORMATION:/note= "RBZ2067-5'"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GGACTCATCA GGTTATCTTT TGTAGTCGTG GG   32

(2) INFORMATION FOR SEQ ID NO: 19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic RNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..23
        (D) OTHER INFORMATION:/note= "Mutant ribozyme 696"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CUGAUGAGUC CCGUGAGGAC GAA   23

(2) INFORMATION FOR SEQ ID NO: 20:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic RNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..21
        (D) OTHER INFORMATION:/note= "Mutant ribozyme 996"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

CUGAUGAGUC CUGAGGACGA A   21

(2) INFORMATION FOR SEQ ID NO: 21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 32 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..32
        (D) OTHER INFORMATION:/note= "RBZ2259-3'"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GTGAGGACGA AAAACCATTC ATAATCACAC CC   32

(2) INFORMATION FOR SEQ ID NO: 22:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 32 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..32
        (D) OTHER INFORMATION:/note= "RBZ 2259-5'"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

GGACTCATCA GATGGTTTGG TGTATCGATA AC   32

(2) INFORMATION FOR SEQ ID NO: 23:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..30
        (D) OTHER INFORMATION:/note= "RBZ2403-3'"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GTGAGGACGA AAAATGGTGC ATGATTTGTC   30

(2) INFORMATION FOR SEQ ID NO: 24:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 28 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..28
        (D) OTHER INFORMATION:/note= "RBZ2403-5'"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GGACTCATCA GTCAGACGCA GCAGAAGC   28

(2) INFORMATION FOR SEQ ID NO: 25:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

        (A) DESCRIPTION: /desc = "synthetic DNA"

   (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..31
        (D) OTHER INFORMATION:/note= "RBZ 2520-3'"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GTGAGGACGA AAAGTCAAAA GCATAGCGTG C   31

(2) INFORMATION FOR SEQ ID NO: 26:

   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 33 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

   (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..33
        (D) OTHER INFORMATION:/note= "RBZ 2520-5'"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

GGACTCATCA GTATGAGGTT ACTTCCAAAA CAC   33

(2) INFORMATION FOR SEQ ID NO: 27:

   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

   (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..30
        (D) OTHER INFORMATION:/note= "RBZ 2720-3'"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GTGAGGACGA AACTGCGGTG GACCCATACC   30

(2) INFORMATION FOR SEQ ID NO: 28:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 32 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..32
            (D) OTHER INFORMATION:/note= "RBZ 2720-5'"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GGACTCATCA GAAGACTAGG TAAACTGGTC AC   32

(2) INFORMATION FOR SEQ ID NO: 29:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 21 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic RNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..21
            (D) OTHER INFORMATION:/note= "Mutant ribozyme 2520"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

CUGAUGAGUC CUGAGGACGA A   21

(2) INFORMATION FOR SEQ ID NO: 30:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 21 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic RNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..21
            (D) OTHER INFORMATION:/note= "Mutant ribozyme 2720"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CUGAUGAGUC GUGAGGACGA A   21

(2) INFORMATION FOR SEQ ID NO: 31:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 20 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
          (A) DESCRIPTION: /desc = "synthetic RNA"

    (ix) FEATURE:
          (A) NAME/KEY: misc_feature
          (B) LOCATION:1..20

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

CUGAUGAGCG GUGACGCGAA   20

(2) INFORMATION FOR SEQ ID NO: 32:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 59 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
          (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
          (A) NAME/KEY: misc_feature
          (B) LOCATION:1..59
          (D) OTHER INFORMATION:/note= "WM2067S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

GCAGCTGCTG GATCCATGTC ATTTTCTTTT CGCGTCACCG CTCATCAGGT TATCTTTTG   59

(2) INFORMATION FOR SEQ ID NO: 33:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 41 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
          (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
          (A) NAME/KEY: misc_feature

        (B) LOCATION:1..41
        (D) OTHER INFORMATION:/note= "WM1747S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

TCAGCTATGT TTCGCGTCAC CGCTCATCAG TGGAGCTTTT C   41

(2) INFORMATION FOR SEQ ID NO: 34:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 49 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..49
        (D) OTHER INFORMATION:/note= "WM996S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

CAGATCTATA AAACTTAGTT TCGCGTCACC GCTCATCAGA TGCCAACTG   49

(2) INFORMATION FOR SEQ ID NO: 35:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 61 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..61
        (D) OTHER INFORMATION:/note= "WM696S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

CAGCAGCTTT TTCGCGTCAC CGCTCATCAG ATGTTCAGGG ACCTGCATCT GATCAAGCTT   60

G   61

(2) INFORMATION FOR SEQ ID NO: 36:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..27
            (D) OTHER INFORMATION:/note= "WMGU4S"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

GAAACATAGC TGACAAAAGA TAACCTG   27

(2) INFORMATION FOR SEQ ID NO: 37:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 26 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..26
            (D) OTHER INFORMATION:/note= "WMGU5S"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

GTTTTATAGA TCTGGAAAAG CTCCAC   26

(2) INFORMATION FOR SEQ ID NO: 38:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 24 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..24
            (D) OTHER INFORMATION:/note= "WMGU6S"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

GAAAAAGCTG CTGCAGTTGG CATC   24

(2) INFORMATION FOR SEQ ID NO: 39:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 55 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..55
        (D) OTHER INFORMATION:/note= "WM2259S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

GCAGCTGCTG GATCCATGAA TGGTTTTTCG CGTCACCGCT CATCAGATGG TTTGG   55

(2) INFORMATION FOR SEQ ID NO: 40:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 42 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..42
        (D) OTHER INFORMATION:/note= "WM2403S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

TGCACCATTT TTCGCGTCAC CGCTCATCAG TCAGACGCAG AT   42

(2) INFORMATION FOR SEQ ID NO: 41:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 41 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..41
        (D) OTHER INFORMATION:/note= "WM2520S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

CTTTTGACTT TTCGCGTCAC CGCTCATCAG TATGAGGTTA C   41

(2) INFORMATION FOR SEQ ID NO: 42:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 61 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..61
        (D) OTHER INFORMATION:/note= "WM2720S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

CACCGCAGTT TCGCGTCACC GCTCATCAGA AGACTAGGTA ACCTGCATCT GATCAAGCTT   60

G   61

(2) INFORMATION FOR SEQ ID NO: 43:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 25 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..25
        (D) OTHER INFORMATION:/note= "WMGU1S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

GAAAAATGGT GCACCAAACC ATCTG   25

(2) INFORMATION FOR SEQ ID NO: 44:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

```
(ii) MOLECULE TYPE: other nucleic acid
     (A) DESCRIPTION:   /desc = "synthetic DNA"

(ix) FEATURE:
     (A) NAME/KEY: misc_feature
     (B) LOCATION:1..27
     (D) OTHER INFORMATION:/note= "WMGU2S"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:
```

CGAAAAGTCA AAAGATCTGC GTCTGAC   27


```
(2) INFORMATION FOR SEQ ID NO: 45:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 25 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
         (A) DESCRIPTION:   /desc = "synthetic DNA"

    (ix) FEATURE:
         (A) NAME/KEY: misc_feature
         (B) LOCATION:1..25
         (D) OTHER INFORMATION:/note= "WMGU3S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:
```

AAACTGCGGT GGTAACCTCA TACTG   25


```
(2) INFORMATION FOR SEQ ID NO: 46:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
         (A) DESCRIPTION:   /desc = "synthetic DNA"

    (ix) FEATURE:
         (A) NAME/KEY: misc_feature
         (B) LOCATION:1..22
         (D) OTHER INFORMATION:/note= "WMPC1S"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:
```

AATTGCAGCT GCTGGATCCA TG   22


```
(2) INFORMATION FOR SEQ ID NO: 47:
```

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 23 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
    (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:
    (A) NAME/KEY: misc_feature
    (B) LOCATION:1..23
    (D) OTHER INFORMATION:/note= "WMPC2S"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

AATTCAAGCT TGATCAGATG CAG   23

(2) INFORMATION FOR SEQ ID NO: 48:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 51 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION:1..51
        (D) OTHER INFORMATION:/note= "FIGURE 2A, WMV2 USA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

GGG AAC AAC AGT GGG CAG CCA TCT ACA GTT GTT GAC AAT ACA CTT ATG   48
GLY ASN ASN SER GLY GLN PRO SER THR VAL VAL ASP ASN THR LEU MET
1   5   10   15

GTT   51
Val

(2) INFORMATION FOR SEQ ID NO: 49:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met

```
 1                  5                  10                 15
Val
```

(2) INFORMATION FOR SEQ ID NO: 50:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 51 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
       (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
       (A) NAME/KEY: CDS
       (B) LOCATION:1..51
       (D) OTHER INFORMATION:/note= "FIGURE 2A, ZYMV"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

```
GGT AAC AAC AGT GGG CAA CCC TCA ACA GTG GTG GAT AAC ACA CTA ATG          48
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
 1               5                   10                  15

GTT                                                                      51
Val
```

(2) INFORMATION FOR SEQ ID NO: 51:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 17 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

```
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
 1               5                   10                  15

Val
```

(2) INFORMATION FOR SEQ ID NO: 52:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 51 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION:1..51
        (D) OTHER INFORMATION:/note= "FIGURE 2A, PLUM POX VIRUS NAT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

```
GGA AAC AAT AGT GGT CAA CCC TCG ACA GTT GTT GAC AAT ACA CTC ATG         48
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
1               5                   10                  15

GTT                                                                     51
Val
```

(2) INFORMATION FOR SEQ ID NO: 53:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 17 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

```
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
1               5                   10                  15

Val
```

(2) INFORMATION FOR SEQ ID NO: 54:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 51 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION:1..51
        (D) OTHER INFORMATION:/note= "FIGURE 2A, TVMV"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

```
GGC AAC AAC AGT GGA CAG CCA TCT ACG GTT GTG GAC AAC ACA CTC ATG         48
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
1               5                   10                  15

GTG                                                                     51
Val
```

(2) INFORMATION FOR SEQ ID NO: 55:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 17 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
1   5   10   15

Val

(2) INFORMATION FOR SEQ ID NO: 56:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 51 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
       (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
       (A) NAME/KEY: CDS
       (B) LOCATION:1..51
       (D) OTHER INFORMATION:/note= "FIGURE 2A, TEV"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

GGC AAC AAT AGC GGG CAA CCT TCA ACA GTG GTG GAC AAC ACA CTC ATG   48
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
1   5   10   15

GTC   51
Val

(2) INFORMATION FOR SEQ ID NO: 57:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 17 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

```
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
1               5               10              15

Val
```

(2) INFORMATION FOR SEQ ID NO: 58:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 51 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION:1..51
        (D) OTHER INFORMATION:/note= "FIGURE 2A, PVY"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

```
GGT AAT AAT AGT GGT CAA CCT TCT ACC GTT GTG GAT AAT TCT CTC ATG        48
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Ser Leu Met
1               5               10              15

GTT                                                                    51
Val
```

(2) INFORMATION FOR SEQ ID NO: 59:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

```
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Ser Leu Met
1               5               10              15

Val
```

(2) INFORMATION FOR SEQ ID NO: 60:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 57 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
     (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:
     (A) NAME/KEY: CDS
     (B) LOCATION:1..57
     (D) OTHER INFORMATION:/note= "FIGURE 2B, PRSV"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

GGC AAT AAT AGT GGC CAA CCT TCG ACA GTT GTT GAT AAT ACA TTG ATG   48
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
1   5   10   15

GTT TTA ATC   57
Val Leu Ile

(2) INFORMATION FOR SEQ ID NO: 61:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 19 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: protein
       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
1   5   10   15

Val Leu Ile

(2) INFORMATION FOR SEQ ID NO: 62:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 57 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (ix) FEATURE:
            (A) NAME/KEY: CDS
            (B) LOCATION:1..57
            (D) OTHER INFORMATION:/note= "FIGURE 2B, PSbMV"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

```
GGA AAT AAT AGT GGA CAA CCA TCA ACA GTT GTT GAT AAC ACA TTG ATG        48
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
 1               5                10               15

GTG ATT CTT                                                            57
Val Ile Leu
```

(2) INFORMATION FOR SEQ ID NO: 63:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 19 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: protein
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

```
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
 1               5                10               15

Val Ile Leu
```

(2) INFORMATION FOR SEQ ID NO: 64:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 57 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (ix) FEATURE:
            (A) NAME/KEY: CDS
            (B) LOCATION:1..57
            (D) OTHER INFORMATION:/note= "FIGURE 2B, TuMV"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

```
GGA AAC AAT AGC GGA CAG CCA TCG ACT GTT GTA GAC AAC ACA CTC ATG        48
Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
 1               5                10               15

GTC ATA TTG                                                            57
Val Ile Leu
```

(2) INFORMATION FOR SEQ ID NO: 65:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 19 amino acids
            (B) TYPE: amino acid

            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

Gly Asn Asn Ser Gly Gln Pro Ser Thr Val Val Asp Asn Thr Leu Met
  1   5   10   15

Val Ile Leu

(2) INFORMATION FOR SEQ ID NO: 66:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (ix) FEATURE:
            (A) NAME/KEY: CDS
            (B) LOCATION:1..30
            (D) OTHER INFORMATION:/note= "FIGURE 3A, WMV2 USA"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

TTC TTT GCC AAT GGT GAT GAT ATT ATC CTG   30
Phe Phe Ala Asn Gly Asp Asp Ile Ile Leu
  1   5   10

(2) INFORMATION FOR SEQ ID NO: 67:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 10 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

Phe Phe Ala Asn Gly Asp Asp Ile Ile Leu
  1   5   10

(2) INFORMATION FOR SEQ ID NO: 68:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid

                    (A) DESCRIPTION: /desc = "synthetic DNA"

            (ix) FEATURE:
                    (A) NAME/KEY: CDS
                    (B) LOCATION:1..30
                    (D) OTHER INFORMATION:/note= "FIGURE 3A, ZYMV"

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

TTC TTT GCA AAT GGA GAT GAT CTG ATA CTT   30
Phe Phe Ala Asn Gly Asp Asp Leu Ile Leu
1   5   10

(2) INFORMATION FOR SEQ ID NO: 69:

        (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 10 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: protein
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

Phe Phe Ala Asn Gly Asp Asp Leu Ile Leu
1   5   10

(2) INFORMATION FOR SEQ ID NO: 70:

            (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 30 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: other nucleic acid
                    (A) DESCRIPTION: /desc = "synthetic DNA"

            (ix) FEATURE:
                    (A) NAME/KEY: CDS
                    (B) LOCATION:1..30
                    (D) OTHER INFORMATION:/note= "FIGURE 3A, PLUM POX VIRUS NAT"

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

TAC TTT GTG AAT GGT GAT GAT CTT GTC CTC   30
Tyr Phe Val Asn Gly Asp Asp Leu Val Leu
1   5   10

(2) INFORMATION FOR SEQ ID NO: 71:

            (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 10 amino acids
                    (B) TYPE: amino acid

```
                    (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: protein
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

    Tyr Phe Val Asn Gly Asp Asp Leu Val Leu
     1               5                   10
```

```
    (2) INFORMATION FOR SEQ ID NO: 72:

            (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 30 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: other nucleic acid
                    (A) DESCRIPTION:   /desc = "SYNTHETIC DNA"

            (ix) FEATURE:
                    (A) NAME/KEY: CDS
                    (B) LOCATION:1..30
                    (D) OTHER INFORMATION:/note= "FIGURE 3A, TVMV"

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

    TTC TTT GCC AAT GGA GAT GAT CTT ATA ATT                    30
    Phe Phe Ala Asn Gly Asp Asp Leu Ile Ile
     1               5                   10
```

```
    (2) INFORMATION FOR SEQ ID NO: 73:

            (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 10 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: protein
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

    Phe Phe Ala Asn Gly Asp Asp Leu Ile Ile
     1               5                   10
```

```
    (2) INFORMATION FOR SEQ ID NO: 74:

            (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 30 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: other nucleic acid
                    (A) DESCRIPTION:   /desc = "synthetic DNA"
```

```
    (ix) FEATURE:
         (A) NAME/KEY: CDS
         (B) LOCATION:1..30
         (D) OTHER INFORMATION:/note= "FIGURE 3A, TEV"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

TAT TAC GTC AAT GGC GAT GAC CTA TTG ATT                          30
Tyr Tyr Val Asn Gly Asp Asp Leu Leu Ile
1               5                   10
```

```
(2) INFORMATION FOR SEQ ID NO: 75:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 10 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

Tyr Tyr Val Asn Gly Asp Asp Leu Leu Ile
1               5                   10
```

```
(2) INFORMATION FOR SEQ ID NO: 76:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION:   /desc = "synthetic DNA"

    (ix) FEATURE:
            (A) NAME/KEY: CDS
            (B) LOCATION:1..30
            (D) OTHER INFORMATION:/note= "FIGURE 3B, PVY"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

TTC TTT GTT AAT GGT GAT GAC TTA TTG ATT                          30
Phe Phe Val Asn Gly Asp Asp Leu Leu Ile
1               5                   10
```

```
(2) INFORMATION FOR SEQ ID NO: 77:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 10 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

Phe Phe Val Asn Gly Asp Asp Leu Leu Ile
 1   5   10

(2) INFORMATION FOR SEQ ID NO: 78:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (ix) FEATURE:
            (A) NAME/KEY: CDS
            (B) LOCATION:1..30
            (D) OTHER INFORMATION:/note= "FIGURE 3B, PRSV"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

TTT TAT ATC AAT GGT GAT GAT CTC TGT ATT   30
Phe Tyr Ile Asn Gly Asp Asp Leu Cys Ile
 1   5   10

(2) INFORMATION FOR SEQ ID NO: 79:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 10 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: protein
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

Phe Tyr Ile Asn Gly Asp Asp Leu Cys Ile
 1   5   10

(2) INFORMATION FOR SEQ ID NO: 80:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (ix) FEATURE:

```
        (A)  NAME/KEY: CDS
        (B)  LOCATION:1..30
        (D)  OTHER INFORMATION:/note= "FIGURE 3B, PSbMV"

    (xi)  SEQUENCE DESCRIPTION: SEQ ID NO: 80:

TAT TTT GCA AAT GGT GAT GAT CTA TTG ATA                              30
Tyr Phe Ala Asn Gly Asp Asp Leu Leu Ile
1               5                   10




(2)  INFORMATION FOR SEQ ID NO: 81:

        (i)  SEQUENCE CHARACTERISTICS:
            (A)  LENGTH: 10 amino acids
            (B)  TYPE: amino acid
            (D)  TOPOLOGY: linear

    (ii)  MOLECULE TYPE: protein
    (xi)  SEQUENCE DESCRIPTION: SEQ ID NO: 81:

Tyr Phe Ala Asn Gly Asp Asp Leu Leu Ile
1               5                   10




(2)  INFORMATION FOR SEQ ID NO: 82:

        (i)  SEQUENCE CHARACTERISTICS:
            (A)  LENGTH: 30 base pairs
            (B)  TYPE: nucleic acid
            (C)  STRANDEDNESS: single
            (D)  TOPOLOGY: linear

    (ii)  MOLECULE TYPE: other nucleic acid
            (A)  DESCRIPTION:   /desc = "synthetic DNA"

    (ix)  FEATURE:
            (A)  NAME/KEY: CDS
            (B)  LOCATION:1..30
            (D)  OTHER INFORMATION:/note= "FIGURE 3B, TuMV"

    (xi)  SEQUENCE DESCRIPTION: SEQ ID NO: 82:

TTC TTC GTC AAC GGA GAC GAT TTA CTG CTA                              30
Phe Phe Val Asn Gly Asp Asp Leu Leu Leu
1               5                   10




(2)  INFORMATION FOR SEQ ID NO: 83:

        (i)  SEQUENCE CHARACTERISTICS:
            (A)  LENGTH: 10 amino acids
            (B)  TYPE: amino acid
            (D)  TOPOLOGY: linear

    (ii)  MOLECULE TYPE: protein
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

Phe Phe Val Asn Gly Asp Asp Leu Leu Leu
1   5   10

(2) INFORMATION FOR SEQ ID NO: 84:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 14 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: Protein
        (B) LOCATION:1..14
        (D) OTHER INFORMATION:/note= "FIGURE 4A, WMV2 USA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

Gly Ser Phe Trp Ile His Trp Ile Thr Thr Gln Asp Gly Phe
1   5   10

(2) INFORMATION FOR SEQ ID NO: 85:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

TTCTGGATAC ATTGGATAAC GACT   24

(2) INFORMATION FOR SEQ ID NO: 86:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 14 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: Protein

          (B) LOCATION:1..14
          (D) OTHER INFORMATION:/note= "FIGURE 4A, Plum Pox Virus NAT"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

Ser His Phe Trp Lys His Trp Ile Ser Thr Lys Asp Gly His
1   5   10

(2) INFORMATION FOR SEQ ID NO: 87:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 24 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: other nucleic acid
          (A) DESCRIPTION: /desc = "synthetic DNA"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

TTCTGGAAAC ACTGGATTAG TACG   24

(2) INFORMATION FOR SEQ ID NO: 88:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 14 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS: unknown
          (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: protein

     (ix) FEATURE:
          (A) NAME/KEY: Protein
          (B) LOCATION:1..14
          (D) OTHER INFORMATION:/note= "FIGURE 4A, TVMV"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

Thr Ser Phe Trp Gln His Trp Ile Thr Thr Lys Asp Gly Gln
1   5   10

(2) INFORMATION FOR SEQ ID NO: 89:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 24 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
          (A) DESCRIPTION: /desc = "synthetic DNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

TTCTGGCAAC ACTGGATAAC AACA   24

(2) INFORMATION FOR SEQ ID NO: 90:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 14 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS: unknown
          (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
          (A) NAME/KEY: Protein
          (B) LOCATION:1..14
          (D) OTHER INFORMATION:/note= "FIGURE 4A, TEV"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

Gly Ile Phe Trp Lys His Trp Ile Gln Thr Lys Asp Gly Gln
1   5   10

(2) INFORMATION FOR SEQ ID NO: 91:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 24 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
          (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

TTCTGGAAGC ATTGGATTCA AACC   24

(2) INFORMATION FOR SEQ ID NO: 92:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 14 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS: unknown
          (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: Protein
        (B) LOCATION:1..14
        (D) OTHER INFORMATION:/note= "FIGURE 4A, PVY"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

Ser Thr Phe Trp Lys His Trp Ile Glu Thr Asp Asn Gly His
1   5   10

(2) INFORMATION FOR SEQ ID NO: 93:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 24 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

TTCTGGAAGC ATTGGATTGA AACA   24

(2) INFORMATION FOR SEQ ID NO: 94:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 14 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: unknown
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: protein

        (ix) FEATURE:
            (A) NAME/KEY: Protein
            (B) LOCATION:1..14
            (D) OTHER INFORMATION:/note= "FIGURE 4A, PRSV"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

Gly Asp Phe Gly Cys His Trp Ile Ser Thr Asn Asp Gly Asp
1   5   10

(2) INFORMATION FOR SEQ ID NO: 95:

        (i) SEQUENCE CHARACTERISTICS:

            (A) LENGTH: 14 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: unknown
            (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
            (A) NAME/KEY: Protein
            (B) LOCATION:1..14
            (D) OTHER INFORMATION:/note= "FIGURE 4B, PSbMV"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

Gly Lys Tyr Trp Lys His Trp Ile Thr Thr Lys Glu Gly His
 1   5   10

(2) INFORMATION FOR SEQ ID NO: 96:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 14 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: unknown
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: protein

        (ix) FEATURE:
            (A) NAME/KEY: Protein
            (B) LOCATION:1..14
            (D) OTHER INFORMATION:/note= "FIGURE 4B, TuMV"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

Ser Gln Phe Trp Lys His Trp Ile Ser Thr Lys Asp Gly Gln
 1   5   10

(2) INFORMATION FOR SEQ ID NO: 97:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 2975 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA to mRNA

        (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..2975
            (D) OTHER INFORMATION:/note= "FIGURES 5(a) to 5(d)"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

```
GAATTCTAGA TTCTGGAAAC ATTGGATAAC CACTCAAGAT GGTTTTTGTG GGCTACCTCT        60

CGTGTCTGTC AATGATGGAT ATATTGTTGG TATACATGGA TTAACATCTA ATGATTCAGA       120

GAAGAATTTC TTTGTTCCAT TTACTGACGG ATTTGAAACG GAATACTTGA ACAATGCTGA       180

TAATCTGTCG TGGGACAAAC ACTGGTTTTG GGAACCAAGT AAGATCGCAT GGGGCTCTTT       240

GAATTTAGTA GAAGAACAGC CAAAGGAAGA ATTCAAGATA TCAAAACTCG TGTCTGATCT       300

TTTTGGCAAC ACGGTAGCAG TACAGAGTAG AAAAGAGAGA TGGGTTTTGG ATGCCATGGA       360

AGGGAATTTG ATAGCTTGTG GGCAAGCTGA TGGCGCACTA GTGACAAAGC ACGTGGTTAA       420

AGGAAAATGC CCTCATTTTG CACAATATCT TTCGTTACAT GATGAAGCAA AACAGTTCTT       480

TGAGCCATTG ATGGGGGCAT ACCAGCCAAG TCGGTTGAAT AAAGACGCTT TCAAGAAAGA       540

TTTCTTTAAG TACAACAAAC CGGTTGTTCT GAATGAAGTT GATTTTAATG CTTTCGAGAA       600

GGCAGTCGAG GGAGTGATAA CAATGATAGT CGACTTTGAA TTTGCAGAAT GCTTATTTGT       660

GACTGATCCT GATGAGATTT ATGGATCCCT GAACATGAAA GCTGCTGTCG GTGCTCAATA       720

CAAAGGAAAG AAACAGGATT ACTTTTCTGG CATGGATAGT TTTGATAAGG AGCGCCTGTT       780

GTACCTTAGC TGTGAGAGGC TTTTTAATGG AGAGAAAGGA ATTTGGAACG GTTCCTTGAA       840

AGCGGAGTTG AGGCCAATTG AAAAAGTGCA GGCAAACAAA ACTAGAACAT TTACAGCAGC       900

ACCGATTGAT ACATTGCTTG GAGCTAAAGT TTGCGTCGAT GATTTAATA ATCAATTCTA        960

CAGCTTCAAC TTAAAATGTC CATGGACAGT TGGCATGACT AAGTTTTATG GGGGTTGGGA      1020

TAAGCTAATG AGGAGTTTGC CTGATGGCTG GACATATTGC CACGCAGACG GTTCACAGTT      1080

CGACAGTTCT TTGACTCCCT TATTGTTGAA TGCTGTCCTC AGTATTAGAT GTTGCTTCAT      1140

GGAAGATTGG TGGGTTGGAA AGGAAATGCT TGAAAACCTT TACGCTGAGA TAGTTTACAC      1200

ACCAATCTTA GCACCTGATG GCACAATTTT CAAAAAGTTT AGAGGGAATA ACAGTGGACA      1260

GCCATCTACA GTTGTTGACA ATACGCTTAT GGTTGTCATT GCTATGTATT ACTCGTGCTG      1320

TAAACAAGGT TGGTCAGAGG AGGATATTGA AGAAAGGTTG GTGTTCTTTG CTAATGGTGA      1380

TGATATTATC TTGGCAGTTA GGGATGAGGA TGTGTGGCTG TATGACACTC TAAGTGCTTC      1440

ATTCGCTGAG CTTGGCCTGA ATTATAACTT CGATGAGCGG ACAAAGAAAA GAGAGGAGCT      1500

ATGGTTCATG TCACACCAAG CCATGCTAGT TGATGGGATT TACATTCCAA AGCTTGAACC      1560

TGAAAGGATA GTTTCCATCT TAGAGTGGGA TAGGAGCAAG GAACTTACGC ATCGCACTGA      1620

GGCAATTTGT GCAGCCATGA TTGAGGCATG GGGACATACT GAGCTTTTGC AGGAAATTCG      1680

CAAATTTTAT TTGTGGCTTC TAAGCAAAGA TGAATTTAAG GAACTCGCTG CATCCGGAAA      1740

AGCTCCATAC ATAGCTGAGA CAGCATTGAG AAAGCTCTAC ACAGATGTTA ACACACAACC      1800

AAGTGAATTG CAAAAATACC TTGAAGTTTT GGATTTTAAC CACATTGATG GTTGCTGTGA      1860

ATCAGTGTCT CTGCAATCAG GAAAAGAAGC AGTAGAAAAT TTGGATGCAG GGAAGGACTC      1920
```

```
GAAGAAGGAC ACCAGTGGCA AAGGGGATAA GCCACAAAAT TCGCAAACTG GTCAAGGTAG        1980

CAAAGAACAG ACAAAAACTG GTACAGTCAG CAAAGATGTG AATGTTGGAT CGAAAGGAAA        2040

AGAAGTCCCA CGACTACAAA AGATAACAAA GAAAATGAAC CTTCCAACAG TTGGTGGGAA        2100

AATCATTCTT AGCTTAGACC ATTTACTCGA GTACAAACCT AATCAAGTTG ATCTGTTTAA        2160

CACTCGAGCA ACAAAAACAC AGTTTGAATC ATGGTACAGC GCAGTTAAAG TTGAATACGA        2220

TCTTAACGAT GAGCAAATGG GTGTGATTAT GAATGGTTTT ATGGTTTGGT GTATCGATAA        2280

CGGTACATCT CCAGATGTCA ATGGAGTTTG GGTGATGATG GATGGGGAAG AGCAAGTTGA        2340

GTATCCATTA AAGCCAATTG TTGAAAATGC GAAACCAACT TTAAGACAAA TCATGCACCA        2400

TTTCTCAGAC GCAGCAGAAG CATATATTGA AATGAGAAAC TCTGAAAGTC CGTATATGCC        2460

TAGATACGGA TTACTAAGAA ATTTGAGAGA CAGGGAATTA GCACGCTATG CTTTTGACTT        2520

TTATGAGGTT ACTTCCAAAG CACCTAATAG GGCAAGAGAA GCAATAGCAC AAATGAAGGC        2580

CGCAGCTCTC GCGGGAATTA ACAGCAGGTT ATTTGGGCTT GATGGTAATA TCTCGACCAA        2640

TTCCGAAAAT ACTGAGAGGC ACACTGCAAG GGACGTGAAT CAGAATATGC ATACTTTGTT        2700

GGGTATGGGT CCACCGCAGT AAAGACTAGG TAAACTGGTC ACAGTTAGCA TTTCGGGTCG        2760

TTATAAGTTT TCTATAATAT AATATGTTGC ACTTTTTAGT ATAGTGTGAT TTTATCACCT        2820

TTGTACTGTT TATGTTAGTG TGGTTTAACC ATCTTAGTGT GCTTTATATT ATAGTTTATG        2880

CATAACAGGG AGAACCATTA CAATACCGGA GTTGTTTGTA GTGTGATTAC ATCACGGTTA        2940

ATAGCCGAGG TACGGTAATG TTTGTTGTCC TAAAA                                 2975
```

(2) INFORMATION FOR SEQ ID NO: 98:
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 2974 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: cDNA to mRNA

       (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..2974
            (D) OTHER INFORMATION:/note= "FIGURE 5 bis (a) to (d)"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

```
GAATTCTAGA TTCTGGAAAC ATTGGATAAG CACTCAAGAT GGTTTTTGTG GGCTACCTCT          60

CGTGTCTGTC AATGATGGAT ATATTGTTGG TATACATGGA TTGACATCTA ATGATTCAGA         120
```

```
GAAGAACTTC TTTGTTCCAT TTACTGACGG ATTTGAAACG GAATACCTGA ACAATGCTGA        180

TAATCTGTCG TGGGACAAAC ACTGGTTTTG GGAACCAAGT AAGATCGCAT GGGGCTCTTT        240

GAATTTAGTA GAAGAACAGC CAAAGGAAGA ATTCAAGATA TCAAAACTCG TGTCTGATCT        300

TTTTGGCAAC ACGGTAGCAG TACAGAGTAG AAAAGAGAGA TGGGTTTTGG ATGCCATGGA        360

AGGGAATTTG ATAGCTTGTG GGCAAGCTGA TAGCGCACTA GTGACAAAGC ACGTGGTTAA        420

AGGAAAATGC CCTCATTTTG CACAATACCT TTCGTTACAT GATGAAGCAA AACAGTTCTT        480

TGAGCCATTG ATGGGGGCAT ACCAGCCAAG TCGGTTGAAT AAAGACGCTT TCAAGAAAGA        540

TTTCTTTAAA TACAACAAAC CGGTTGTTCT GAATGAAGTT GATTTAATG CTTTCGAGAA         600

GGCAGTCGAG GGAGTGGTAA CAATGATGGT CGACTTTGAA TTTGCAGAAT GCTTATTTGT        660

GACTGATCCT GATGAGATTT ATGGATCCCT GAACATGAAA GCTGCTGTCG GTGCTCAATA        720

CAAAGGAAAG AAACAGGATT ACTTTTCTGG CATGGATAGT TTTGATAAGG AGCGCCTGTT        780

GTACCTTAGC TGTGAGAGGC TTTTTAACGG GGAGAAAGGA ATTTGGAATG GTTCCTTGAA        840

AGCGGAGTTG AGGCCAATTG AAAAAGTGCA GGCGAACAAA ACTAGAACAT TTACAGCAGC        900

ACCAATTGAC ACATTGCTTG GAGCTAAAGT TTGCGTCGAT GATTTAATA ATCAATTCTA         960

TAGCTTCAAC TTAAAATGTC CATGGACAGT TGGCATGACT AAGTTTTATG GGGGTTGGGA       1020

TAAGCTAATG AGGAGTTTGC CTGATGGCTG GACATATTGC CACGCAGACG GTTCACAGTT       1080

CGACAGTTCT TTGACTCCCT TATTGTTGAA TGCTGTCCTC AGTATTAGAT GTTGCTTCAT       1140

GGAAGATTGG TGGGTCGGAA AGGAAATGCT CGAAAACCTT TACGCTGAGA TAGTTTACAC       1200

ACCAATCTTA GCACCTGATG GCACAATTTT CAAAAAGTTT AGAGGGAATA ACAGTGGACA       1260

GCCATCTACA GTTGTTGACA ATACGCTTAT GGTTGTCATT GCTATGTATT ACTCGTGCTG       1320

TAAGCAAGGT TGGTCAGAGA AGGATATTGA AGAAAGGTTG GTGTTCTTTG CTAATGGTGA       1380

TGATATTATC TTGGCAGTTA GGGACGAGGA TGTGTGGCTG TATGACACTC TAAGTGCTTC       1440

ATTCGCTGAG CTTGGCCTGA ATTATAACTT CGATGAGCGG ACAAAGAAAA GAGAGGAGCT       1500

ATGGTTTATG TCACACCAAG CCATGCTAGT TGATGGGATT TATATTCCAA AGCTTGAACC       1560

TGAAAGGATA GTTTCCATCT TAGAGTGGGA TAGGAGCAAG GAACTTATGC ATCGCACTGA       1620

GGCAATTTGT GCAGCCATGA TTGAGGCATG GGACATACT GAGCTTTTGC AGGAAATTCG        1680

CAAATTTTAT TTGTGGCTTC TAAGCAAAGA TGAATTTAAG GAACTCGCTG CATCCGGAAA       1740

AGCTCCATAC ATAGCTGAGA CAGCATTGAG AAAGCTCTAC ACAGATGTTA ACACACAACC       1800

AAGTGAATTG CAAAAATACC TTGAAGTTTT GGATTTTAAC CACATTGATG GTTGTTGTGA       1860

ATCAGTGTCT CTGCAATCAG GAAAAGAAGC AGTAGAAAAT TTGGATGCAG GGAAGGACTC       1920

GAAGAAGGAC ACCAGTGGCA AAGGGGATAA GCCACAAAAC TCGCAAACTG ATCAAGGTAG       1980

CAAAGAACAG ACAAAAACTG GCACAGTCAG CAAAGATGTG AATGTTGGAT CGAAAGGAAA       2040
```

```
AGAAGTCCCA CGACTACAAA AGATAACAAA GAAAATGAAC CTTCCAACAG TTGGTGGGAA        2100

AATCATTCTT AGCTTAGACC ATTTGCTCGA GTACAAACCT AATCAAGTTG ATCTGTTTAA        2160

CACTCGAGCA ACAAAAACAC AGTTTGAGTC ATGGTACAGC GCAGTTAAAG TTGAATATGA        2220

TCTTAACGAT GAGCAAATGG GTGTGATTAT GAATGGTTTT ATGGTTTGGT GTATCGATAA        2280

CGGTACATCT CCAGATGTCA ATGGAGTTTG GGTGATGATG GATGGGGAAG AGCAAGTTGA        2340

GTATCCATTA AAGCCAATTG TTGAAAATGC AAAACCAACT TTAAGACAAA TCATGCACCA        2400

TTTCTCAGAC GCAGCAGAAG CATATATTGA AATGAGAAAC TCTGAAAGTC CGTATATGCC        2460

TAGATACGGA TTACTAAGAA ATTTGAGAGA CAGGGAATTA GCACGCTATG CTTTTGACTT        2520

TTATGAGGTT ACTTCCAAAG CACCTAATAG GGCAAGAGAA GCAATAGCAC AAATGAAGGC        2580

CGCAGCTCTC GCGGGAATTA ACAGCAGGTT ATTTGGGCTT GATGGTAATA TCTCGACCAA        2640

TTCCGAAAAT ACTGAGAGGC ACACTGCAAG GGACGTGAAT CAGAATATGC ATACTTTGTT        2700

AGGTATGGGT CCACCGCAGT AAAGACTAGG TAAACTGGTC ACAGTTAGCA TTTCGGGTCG        2760

TTACAAGTTT TCTATAATAT AATATGTTGC ACTTTTTAGT ATAGTGTGAT TTTATCACCT        2820

TTGTACTGTT TATGTAGTGT GGTTTAACCA CCTTAGTGTG CTTTATATTA TAGTTTATGC        2880

ATAACAGGGA GAACCATTGC AATACCGGAG TTGTTTGTAG TGTGATTACA TCACGTTTAA        2940

TAGCCGAGGT ACGGTAATGT TTGTTGTCCT AAAA                                    2974
```

(2) INFORMATION FOR SEQ ID NO: 99:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2922 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA to mRNA

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..2922
        (D) OTHER INFORMATION:/note= "FIGURE 6(a) to (d)"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

```
TCTAGATTCT GGAAACATTG GATAACACCA ACGAAGGGGA TTGCGGATTG CCCATGGTTT          60

CAACAACGGA TGGCAAGATA ATTGGAGTTC ATGGTTTGGC TTCCACAGTC TCATCTAAGA         120

ATTATTTTGT CCCATTCACT GATGATTTTA TAGCCACGCA TTTGAGCAAA CTTGATGACC         180

TCACATGGAC TCAGCATTGG CTATGGCAAC CTAGCAAAAT TGCGTGGGGA ACGCTCAACT         240

TAGTTGATGA ACAACCAGGG CCCGAATTTC GTATTTCAAA TCTAGTCAAG GATTTATTCA         300
```

```
CTTCTGGTGT TGAAACACAG AGCAAGCGAG AAAGATGGGT CTACGAAAGC TGTGAAGGGA          360

ACCTTCGGGC TGTTGGAACT GCACAATCAG CGTTAGTCAC CAAACATGTT GTGAAAGGCA          420

AGTGTCCTTT CTTCGAAGAA TATTTACAAA AACACGCAGA AGCGAGCGCC TATTTCAGAC          480

CCCTAATGGG AGAGTACCAG CCGAGCAAGT TGAACAAAGA AGCCTTTAAA AAGGATTTCT          540

TTAAATACAA TAAACCCGTC ACTGTTAATC AACTGGATCA TGATAAATTT TTGGAAGCAG          600

TGGATGGGGT TATACGTATG ATGTGTGACT TTGAGTTCAA CGAATGTCGA TTCATTACAG          660

ATCCCGAGGA GATTTACAAC TCTTTGAACA TGAAAGCAGC AATTGGAGCC CAGTATAGAG          720

GAAAGAAGAA AGAGTTTTTG AGGGGCTAGA TGATTTTGAT CGAGAGCGAC TTTTATTCCA          780

AAGTTGTGAA AGGTTGTTCA ATGGCTACAA AGGTCTGTGG AATGGATCTT TAAAGGCCGA          840

GCTCAGGCCG CTTGAGAAAG TCAGGGCTAA CAAAACACGA ACCTTTACAG CAGCGCCAAT          900

TGATACATTG CTTGGAGCTA AAGTTTGTGT GGATGATTTC AACAATGAGT TCTACAGGAA          960

AAACCTCAAG TGTCCATGGA CGGTCGGCAT GACAAAATTT TATGGTAGTT GGGATAAATT         1020

GATGAGATCA TTACCTGATG GTTGGTTGTA TTGTCATGCT GATGGATCAC AGTTCGATAG         1080

TTCGTTAACC CCAGCCTTAC TGAACGCAGT GCTCATAATC AGGTCATTTT ATATGGAGGA         1140

TTGGTGGGTT GGCCAAGAGA TGCTTGAAAA TCTTTATGCC GAGATTGTGT ACACTCCAAT         1200

TCTTGCTCCT GATGGAACAA TTTTCAAGAA ATTTAGAGGT AACAACAGTG GGCAACCCTC         1260

AACAGTGGTG GATAACACAC TAATGGTTGT GATCTCTATT TACTATGCGT GCATGAAATT         1320

TGGTTGGAAC TGCGAGGAGA TTGAGAATAA ACTTGTCTTC TTTGCAAATG GGGATGATCT         1380

GATACTTGCA GTCAAAGATG AGGATAGCGG CTTACTTGAT AACATGTCAT CCTCTTTTTG         1440

CGAACTTGGA CTGAATTATG ATTTTTCAGA ACGTACGCAT AAAAGAGAAG ATCTTTGGTT         1500

CATGTCCCAC CAAGCAATGC TAGTTGATGG AATGTACATT CCAAAACTCG AGAAAGAGAG         1560

AATTGTTTCA ATTCTAGAGT GGGATAGAAG CAAAGAAATT ATGCACCGAA CAGAGGCTAT         1620

TTGCGCTGCG ATGATTGAGG CATGGGGGCA CACCGAGCTC TTGCAAGAAA TCAGAAAGTT         1680

TTACCTATGG TTCGTTGAAA AAGAAGAAGT GCGAGAATTG GCAGCCATCG GGAAAGCTCC         1740

ATACATAGCT GAGACAGCAC TTCGTAAGTT ATACACTGAC AAGGGAGCAG ATACAAGTGA         1800

ACTGGCACGC TACCTACAAG CCCTCCATCA AGACATCTTC TTTGAACAAG AGACACTGT          1860

GATGCTCCAA TCAGGCACTC AGCCAACTGT GGCAGATGCT GGAGCTACAA AGAAAGACAA         1920

AGAAGATGAC AAAGGGAAAA ACAAGGACGT TACAGGCTCT GGCTCAGGTG AGAAAACAGT         1980

AGCAGCTGTC ACGAAGGACA AGGATGTGAA TGCTGGTTCT CATGGGAAAA TTGTGCCGCG         2040

TCTTTCGAAG ATCACAAAGA AAATGTCATT GCCACGCGTG AAAGGAAATG TGATACTCGA         2100

TATTGATCAT TTGCTGGAAT ATAAACCGGA TCAAATTGAG TTATATAACA CACGAGCGTC         2160
```

```
TCATCAGCAG TTCGCCTCTT GGTTCAACCA GGTTAAGACG GAATATGATT TGAACGAGCA        2220

ACAGATGGGA GTTGTAATGA ATGGTTTCAT GGTTTGGTGT ATTGAAAATG GCACTTCACC        2280

CGACATTAAT GGAGTGTGGG TTATGATGGA CGGAAATGAG CAAGTTGAGT ATCCCTTGAA        2340

ACCAATAGTT GAAAATGCAA AGCCAACGCT GCGGCAAATA ATGCATCATT TTTCAGATGC        2400

AGCGGAGGCA TATATAGAGA TGAGAAATGC AGAGGCACCA TACATGCCGA GGTATGGTTT        2460

GCTTCGAAAC CTACGGGATA GGAGTTTAGC ACGATATGCT TTCGATTTCT ATGAAGTCAA        2520

TTCTAAAACT CCTGAAAGAG CCCGCGAAGC TGTTGCGCAG ATGAAAGCAG CAGCTCTTAG        2580

CAATGTTTCT TCAAGGTTGT TTGCCTTGAT GGAAATGTTG CCACCACTAG TGAAGACACT        2640

GAACGGCACA CTGCACGTGA TGTTAATAGA AACATGCACA CCTTACTAGG TGTGAATACA        2700

ATGCAGTAAA GGGTAGGCCG CCTACCTAGG TTATTGTTTC GCTGCCGACG TAATTCTAAT        2760

ATTTACCGCT TTATTTGATA TCTTTAGATT TCCAGAGTGG GCCTCCCACC TTTAAAGCGT        2820

AAAGTTTATG TTAGTTGTCC AGGAGTGCCG TAGTCCTGTC GGAAGCTTTA GTGTGAGCCT        2880

CTCACGAATA AGCTCGAGAT TAGACTCCGT TTGAAGCCTA AA                           2922
```

(2) INFORMATION FOR SEQ ID NO: 100:
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 236 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..236
        (D) OTHER INFORMATION:/note= "FIGURE 7A"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

```
GAGCTCCACC GCGGTGGCGG CCGCTCTAGA ACTAGTGGAT CCATGTCATT TTCTTTTCGC         60

GTCACCGCTC ATCAGGTTAT CTTTTGTCAG CTATGTTTCG CGTCACCGCT CATCAGTGGA        120

GCTTTTCCAG ATCTATAAAA CTTAGTTTCG CGTCACCGCT CATCAGATGC CAACTGCAGC        180

AGCTTTTTCG CGTCACCGCT CATCAGATGT TCAGGGACCT GCATCTGATC AAGCTT           236
```

(2) INFORMATION FOR SEQ ID NO: 101:

        (i) SEQUENCE CHARACTERISTICS:

```
            (A) LENGTH: 230 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION:   /desc = "synthetic DNA"

      (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..230
            (D) OTHER INFORMATION:/note= "FIGURE 7B"

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:
```

GAGCTCCACC GCGGTGGCGG CCGCTCTAGA ACTAGTGGAT CCATGAATGG TTTTTCGCGT   60

CACCGCTCAT CAGATGGTTT GGTGCACCAT TTTTCGCGTC ACCGCTCATC AGTCAGACGC   120

AGATCTTTTG ACTTTTCGCG TCACCGCTCA TCAGTATGAG GTTACCACCG CAGTTTCGCG   180

TCACCGCTCA TCAGAAGACT AGGTAACCTG CATCTGATCA AGCTTGAATT   230


```
  (2) INFORMATION FOR SEQ ID NO: 102:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 348 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION:   /desc = "synthetic DNA"

      (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..348
            (D) OTHER INFORMATION:/note= "FIGURE 7C"

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:
```

AAGCTTGATA TCGAATCAGA ATATGCATAC TTTGTTAGGT ATGGGTCCAC CGCAGTAAAG   60

ACTAGGTAAA CTGGTCACAG TTAGCATTTC GGGTCGTTAC AAGTTTTCTA TAATATAATA   120

TGTTGCACTT TTTAGTATAG TGTGATTTTA TCACCTTTGT ACTGTTTATG TAGTGTGGTT   180

TAACCACCTT AGTGTGCTTT ATATTATAGT TTATGCATAA CAGGGAGAAC CATTGCAATA   240

CCGGAGTTGT TTGTAGTGTG ATTACATCAC GTTTAATAGC CGAGGTACGG TAATGTTTGT   300

TGTCCTAAAA AAAAAAGGAT CCACTAGTTC TAGAGCGGCC GCCACCGC   348


```
  (2) INFORMATION FOR SEQ ID NO: 103:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 227 base pairs
```

```
              (B)  TYPE: nucleic acid
              (C)  STRANDEDNESS: single
              (D)  TOPOLOGY: linear

        (ii)  MOLECULE TYPE: other nucleic acid
              (A)  DESCRIPTION:    /desc = "SYNTHETIC DNA"

        (ix)  FEATURE:
              (A)  NAME/KEY: misc_feature
              (B)  LOCATION:1..227
              (D)  OTHER INFORMATION:/note= "FIGURE 8A"

        (xi)  SEQUENCE DESCRIPTION: SEQ ID NO: 103:
```

GCGGCCGCTC TAGAACTAGT GGATCCATGC AATGTTTTTC GCGTCACCGC TCATCAGTTC   60

AAGGTGCCGA GGTTTCGCGT CACCGCTCAT CAGTGGTTTG CTTCATGGTT TCGCGTCACC   120

GCTCATCAGT GGTGCATACG AGCGTTTCGC GTCACCGCTC ATCAGTCATC AGCTCCACAG   180

TTTCGCGTCA CCGCTCATCA GTCATCTAAC TGCATCTGAT CAAGCTT   227

```
  (2)  INFORMATION FOR SEQ ID NO: 104:

        (i)  SEQUENCE CHARACTERISTICS:
              (A)  LENGTH: 226 base pairs
              (B)  TYPE: nucleic acid
              (C)  STRANDEDNESS: single
              (D)  TOPOLOGY: linear

        (ii)  MOLECULE TYPE: other nucleic acid
              (A)  DESCRIPTION:    /desc = "synthetic DNA"

        (ix)  FEATURE:
              (A)  NAME/KEY: misc_feature
              (B)  LOCATION:1..226
              (D)  OTHER INFORMATION:/note= "FIGURE 8B"

        (xi)  SEQUENCE DESCRIPTION: SEQ ID NO: 104:
```

AAGCTTGATC AGATGCAGGC TGCTTTTTCG CGTCACCGCT CATCAGATCT GCGCCGGCAT   60

GTTTCGCGTC ACCGCTCATC AGTGGTGCCT CCCTTTGTTT CGCGTCACCG CTCATCAGAT   120

CTTCTTAGCT ATGTTTCGCG TCACCGCTCA TCAGTGGAGC TTCCCACTGT TTCGCGTCAC   180

CGCTCATCAG GTTACCTCAT GGATCCACTA GTTCTAGAGC GGCCGC   226

```
  (2)  INFORMATION FOR SEQ ID NO: 105:

        (i)  SEQUENCE CHARACTERISTICS:
              (A)  LENGTH: 220 base pairs
              (B)  TYPE: nucleic acid
              (C)  STRANDEDNESS: single
              (D)  TOPOLOGY: linear

        (ii)  MOLECULE TYPE: other nucleic acid
```

```
            (A) DESCRIPTION:    /desc = "synthetic DNA"

      (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..220
            (D) OTHER INFORMATION:/note= "FIGURE 8C"

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:
```

GGATCCATGC TTTGCATTTT CGCGTCACCG CTCATCAGTC AACTATCGAT GCCGAGGTTT   60

CGTCCTCACG GACTCATCAG TGGTTTGCTT TTCGTCCTCA CGGACTCATC AGGAAACCTA   120

CGCTGCAGCA ATGTTTTTCG TCCTCACGGA CTCATCAGTT CAAGGTTGTT TTCGCGTCAC   180

CGCTCATCAG GGCCTTGATG ACCTGCATCT GATCAAGCTT   220

```
(2) INFORMATION FOR SEQ ID NO: 106:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 93 base pairs
           (B) TYPE: nucleic acid
           (C) STRANDEDNESS: single
           (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: other nucleic acid
           (A) DESCRIPTION:    /desc = "synthetic DNA"

      (ix) FEATURE:
           (A) NAME/KEY: misc_feature
           (B) LOCATION:1..93
           (D) OTHER INFORMATION:/note= "FIGURE 8D"

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:
```

GGATCCGCAC AATTTTTTCG CGTCACCGCT CATCAGCCAT GAGAACCAGC ATTTTCGCGT   60

CACCGCTCAT CAGACATCCT TGTGATCAAG CTT   93

```
(2) INFORMATION FOR SEQ ID NO: 107:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 1054 base pairs
           (B) TYPE: nucleic acid
           (C) STRANDEDNESS: single
           (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: other nucleic acid
           (A) DESCRIPTION:    /desc = "synthetic DNA"

      (ix) FEATURE:
           (A) NAME/KEY: misc_feature
           (B) LOCATION:1..1054
           (D) OTHER INFORMATION:/note= "FIGURE 9(a) and (b)"

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:
```

```
TTATTTGCCG CAGCGTTGGC TTTGCATTTT CGTCCTCACG GACTCATCAG TCAACTATTG          60

GTTTCAAGGG ATACTCAACT TGCTCATTTC CGTCCATCAT AACCCACACT CCATTAATGT         120

CGGGTGAAGT GCCATTTTCA ATACACCAAA CCATGAAACC ATTCATTACA ACTCCCATCT         180

GTTGCTCGTT CAAATCATAT TCCGTCTTAA CCTGGTTGAA CCAAGAGGCG AACTGCTGAT         240

GAGACGCTCG TGTGTTATAT AACTCAATTT GATCCGGTTT ATATTCCAGC AAATGATCAA         300

TATCGAGTAT CACATTTCCT TTCACGCGTG GCAATGACAT TTTCTTTGTG ATCTTCGAAA         360

GACGCGGCAC AATTTTTTCG TCCTCACGGA CTCATCAGCC ATGAGAACCA GCATTTTCGT         420

CCTCACGGAC TCATCAGACA TCCTTGTCCT TCGTGACAGC TGCTACTGTT TTCTCACCTG         480

AGCCAGTCAT TTTTCAGATG CAGCGGAGGC ATATATAGAG ATGAGAAATG CAGAGGCACC         540

ATACATGCCG AGGTTTCGTC CTCACGGACT CATCAGTGGT TTGCTTTTCG TCCTCACGGA         600

CTCATCAGGA AACCTACGGG ATAGGAGTTT AGCACGATAT GCTTTCGATT TCTATGAAGT         660

CAATTCTAAA ACTCCTGAAA GAGCCCGCGA AGCTGTTGCG CAGATGAAAG CAGCAGCTCT         720

TAGCAATGTT TTTCGTCCTC ACGGACTCAT CAGTTCAAGG TTGTTTTCGT CCTCACGGAC         780

TCATCAGGGC CTTGATGGAA ATGTTGCCAC CACTAGTGAA GACACTGAAC GGCACACTGC         840

ACGTGATGTT AATAGAAACA TGCACACCTT ACTAGGTGTG AATACAATGC AGTAAAGGGT         900

AGGCCGCCTA CCTAGGTTAT TGTTTCGCTG CCGACGTAAT TCTAATATTT ACCGCTTTAT         960

TTGATATCTT TAGATTTCCA GAGTGGGCCT CCCACCTTTA AAGCGTAAAG TTTATGTTAG        1020

TTGTCCAGGA GTGCCGTAGT CCTGTCGGAA GCTT                                    1054
```

(2) INFORMATION FOR SEQ ID NO: 108:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 239 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION:1..239
        (D) OTHER INFORMATION:/note= "FIGURE 18(-)"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

```
GTTTAAAATA AACACCGAAG ATTCGTTTCT ACTTAAATTC CTTGAGCGAC GTAGGCCTTT          60

TCGAGGTATG TATCGACTCT GTCGTAACTC TTTCGAGATG TGTCTACAAT TGTGTGTTGG         120

TTCACTTAAC GTTTTTATGG AACTTCAAAA CCTAAAATTG GTGTAACTAC CAACGACACT         180
```

TATCACAGAG ACGTTAGTCC TTTTCTTCGT CATCTTTTAA ACCTACGTCC CTTCCTGAG   239

(2) INFORMATION FOR SEQ ID NO: 109:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 2005 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION:1..2005
            (D) OTHER INFORMATION:/note= "FIGURE 19"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

TATCAAAGCT CGTGTCTGAT CTTTTTGGCA ACACGGTAGC AGTACAGAGT AGAAAAGAGA   60

GATGGGTTTT GGACGCCATG GAAGGGAATT TGATAGCTTG TGGGCAAGCT GATAGCGCAC   120

TAGTGACAAA GCACGTGGTT AAAGGAAAAT GCCCTCATTT TGCACAATAT CTTTCGTTAC   180

ATGATGAAGC AAAACAGTTC TCTGAGCCAT TGATGGGGGC ATACCAGCCA AGTCGGTTGA   240

ATAAAGACGC TTTCAAGAAA GATTTCTTTA AGTACAACAA ACCGGTTGTT CTGAATGAAG   300

TTGATTTTAA TGCTCTCGAG AAGGCAGTCG AGGGAGTGAT AACAATGATG GTCGACTTTG   360

AATTTGCAGA ATGCTTATTT GTGACTGATC CTGATGAGAT TTATGGATCC CTGAACATCT   420

GATGAGTCCC GTGAGGACGA AAAAGCTGCT GTCGGTGCTC AATACAAAGG AAAGAAACAG   480

GATTACTTTT CTGGCATGGA TAGTTTTGAT AAGGAGCGCC TGTTGTACCT TAGCTGTGAG   540

AGGCTTTTTA ATGGAGAGAA AGGAATTTGG AACGGTTCCT TGAAAGCGGA GTTGAGGCCA   600

ATTGAAAAAG TGCAGGCAAA CAAAACTAGA ACATTTACAG CAGCACCGAT TGATACATTG   660

CTTGGAGCTA AAGTTTGCGT CGATGATTTT AATAATCAAT TCTACAGCTT CAACTTAAAA   720

TGTCCATGGA CAGTTGGATC TGATGAGTCC TGAGGACGAA ACTAAGTTTT ATGGGGGTTG   780

GGATAAGCTA ATGAGGAGTT TTGCCTGATG GCTGGACATA TTGCCACGCA GACGGTTCAC   840

AGTTCGACAG TTCTTTGACT CCCTTATTGT TGAATGCTGT CCTCAGTATT AGATGTTGCT   900

TCATGGAAGA TTGGTGGGTT GGAAAGGAAA TGCTTGAAAA CCTTTACGCT GAGATAGTTT   960

ACACACCAAT CTTAGCACCT GATGGCACAA TTTTCAAAAA GTTTAGAGGG AATAACAGTG   1020

GACAGCCATC TACAGTTGTT GACAATACGC TTATGGTTGT CATTGCTATG TATTACTCGT   1080

GCTGTAAACA AGGTTGGTCA GAGGAGGATA TTGAAGAAAG GTTGGTGTTC TTTGCTAATG   1140

```
GTGATGATAT TATCTTGGCA GTTAGGGATG AGGATGTGTG GCTGTATGAC ACTCTAAGTG          1200

CTTCATTCGC TGAGCTTGGC CTGAATTATA ACTTCGATGA GCGGACAAAG AAAAGAGAGG          1260

AGCTATGGTT CATGTCACAC CAAGCCATGC TAGTTGATGG GATTTACATT CCAAAGCTTG          1320

AACCTGAAAG GATAGTTTCC ATCTTAGAGT GGGATTGGAG CAAGGAACTT ATGCATCGCA          1380

CTGAGGCAAT TTGTGCAGCC ATGATTGAGG CATGGGGACA TACTGAGCTT TTGCAGGAAA          1440

TTCGCAAATT TTATTTGTGG CTTCTAAGCA AAGATGAATT TTAAGGAACT CGCTGCATCC          1500

GGAAAAGCTC CACTGATGAG TCCGTGAGGA CGAAACATAG CTGAGACAGC ATTGAGAAAG          1560

CTCTACACAG ATGTTAACAC ACAACCAAGT GAATTGCAAA AATACCTTGA AGTTTTGGAT          1620

TTTAACCACA TTGATGGTTG CTGTGAATCA GTGTCTCTGC AATCAGGAAA AGAAGCAGTA          1680

GAAAATTTGG ATGCAGGGAA GGACTCGAAG AAGGACACCA GTGGCAAAGG GGATAAGCCA          1740

CAAAATTCGC AAACTGGTCA AGGTAGCAAA GAACAGACAA AAACTGGTAC AGTCAGCAAA          1800

GATGTGAATG TTGGATCGAA AGGAAAAGAA GTCTCACGAC TACAAAAGAT AACCTGATGA          1860

GTCCGTGAGG ACGAAAAGAA AATGAACCTT CCAACAGTTG GTGGGAAAAT CATTCTTAGC          1920

TTAGCCATTT ACTCGAGTAC AAACCTAATC AAGTTGATCT TTAACACTCG AGCCCAGCTT          1980

CCACCATGGC GTGCAGGTCG ACGGA                                              2005
```